Random sections of ellipsoids and the power of random information

joint work with

David Krieg, Erich Novak, Joscha Prochno, Mario Ullrich

Aicke Hinrichs

Johannes Kepler University Linz

Cambridge, Isaac Newton Institute, February 2019

Introduction

Problems and Algorithms - IBC Setting

- **Numerical Problem:** linear solution operator $S : F \to G$
- normed space F of inputs
- normed space G of outputs, norm measures error
- **Examples:** Integration, Approximation
- **Information:** $N_n : F \to \mathbb{R}^n$ linear
- maybe restricted to a subclass like n function evaluations
- **Algorithm:** $A_n = \varphi \circ N_n$ with $\varphi : \mathbb{R}^n \to G$

Error and Radius of Information

- **Worst Case Error:**

  $e(S, A_n) = \sup\{\|Sf - \varphi(N_n f)\|_G : \|f\|_F \le 1\}$

- **Radius of Information:** $r(S, N_n) = \inf_{\varphi} e(S, \varphi \circ N_n)$

- **Radius at Zero:**

  $r_0(S, N_n) = \sup\{\|Sf\|_G : \|f\|_F \le 1, N_n f = 0\}$

- **Result:** $r_0(S, N_n) \le r(S, N_n) \le 2r_0(S, N_n)$

- **Minimal Worst Case Error:**

  $e_n(S) = \inf_{\varphi, N_n} e(S, \varphi \circ N_n) = \inf_{N_n} r(S, N_n)$

- **Optimal Information:** N_n^* with $r(S, N_n^*) = e_n(S)$

Random Information vs. Optimal Information

- **Random Information:** $N_n f = (L_1(f), \dots, L_n(f))$ where the linear functionals $L_1, \dots, L_n$ are taken as i.i.d. random variables
- **Expected Radius of Information** $\mathbb{E}\, r(S, N_n)$
- **Clearly**
$$r(S, N_n^*) = \inf_{N_n} r(S, N_n) \le \mathbb{E}\, r(S, N_n)$$
- **General Question:** How good is random information compared to optimal information, i.e., how do $r(S, N_n^*)$ and $\mathbb{E}\, r(S, N_n)$ compare?

Example

Take $S = Id : F \to G$ with $G = L_p$ and

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

Information: function values, $N_n(f) = (f(x_1), \dots, f(x_n))$.

The optimal info does not depend on p,

$N_n^*(f) = (f(\frac{1}{2n}), f(\frac{3}{2n}), \dots, f(\frac{2n-1}{2n}))$.

Moreover $e_n(S)$ is between $\dfrac{1}{2n}$ (for $p = \infty$) and $\dfrac{1}{4n}$ (for $p = 1$).

Now take i.i.d. info with unif. distr. x_i.

Then $\mathbb{E}\, r(S, N_n)$ is not much larger then $r(S, N_n^*)$:

- $\mathbb{E}\, r(S, N_n) = \dfrac{n+3}{2(n+1)(n+2)} \approx \dfrac{1}{2n}$ for $p = 1$
- $\mathbb{E}\, r(S, N_n) \asymp \dfrac{1}{n}$ for $p < \infty$
- $\mathbb{E}\, r(S, N_n) \asymp \dfrac{\log n}{n}$ for $p = \infty$.

Hilbert Space Approximation Problem

The Problem

We study random information for L_2-approximation of functions from a Hilbert space.

- $\sigma_1 \geq \sigma_2 \geq \ldots \geq 0$
- $m \in \mathbb{N} \cup \{\infty\}$
- $F = \left\{ x \in \mathbb{R}^m \,:\, \|x\|_F^2 = \sum_{j=1}^{m} \frac{x_j^2}{\sigma_j^2} < \infty \right\}$
- $G = \ell_2^m = \left\{ x \in \mathbb{R}^m \,:\, \|x\|_{\ell_2^m}^2 = \sum_{j=1}^{m} x_j^2 < \infty \right\}$
- $S = Id : F \to G$

Optimal Information ...

- is given by $N_n^* x = (x_1, \dots, x_n)$ ...
- ... and gives the minimal worst case error aka minimal radius of information

$$e_n(S) = r(S, N_n^*) = \sigma_{n+1}.$$

Random Information

- We want to compare this to random Gaussian information given by $L_j x = \sum_{i=1}^m g_{ij} x_i$ ...
- ... where g_{ij} are i.i.d. standard Gaussian random variables
- equivalently for finite m: $L_j x = \langle x, y_j \rangle$ where $y_1, \ldots, y_n$ are i.i.d. uniform on the Euclidean unit sphere of $\mathbb{R}^m$
- **Geometric Formulation** for finite m: What is the expected radius of an ellipsoid obtained by slicing the unit ellipsoid of F with an n-codimensional random linear subspace of $\mathbb{R}^m$ (with respect to the uniform distribution on the Grassmannian manifold)?

$m = \infty$ - Do we have bounded information?

- Problem for $m = \infty$: $L_j x = \sum_{i=1}^m g_{ij} x_i$ can be infinite.
- But: If $\sigma \in \ell_2$ then $L_j : F \to \mathbb{R}$ is finite almost surely for $j = 1, \ldots, n$ and

$$\mathbb{E}\, \| N_n : F \to \ell_2^n \|^2 \leq n \|\sigma\|_{\ell_2}^2 .$$

- On the other hand: If $\sigma \notin \ell_2$ then the N_n-image of the finite sequences in the unit ball of F already is (almost surely) the whole $\mathbb{R}^n$.
- We formulate the results mainly for $m = \infty$.

Results

The $\sigma \in \ell_2$ – dichotomy

- Random information behaves very differently depending on whether $\sigma \in \ell_2$ or not.
- For $\sigma \notin \ell_2$ random information is completely useless, i.e., $\mathbb{E}\, r(S, N_n) = \sigma_1$.
- For $\sigma \in \ell_2$ the expected radius of random information $\mathbb{E}\, r(S, N_n)$ tends to zero at least at rate $o(1/\sqrt{n})$ as $n \to \infty$.

The $\sigma \in \ell_2$ – dichotomy - closer look

- Consider the case

$$\sigma_k \asymp k^{-\alpha} \ln^{-\beta}(k+1),$$

where $\alpha > 0$ and $\beta \in \mathbb{R}$

- Then

$$\mathbb{E}\, r(S, N_n) \asymp \begin{cases} \sigma_1 & \text{if } \alpha < 1/2 \text{ or } \beta \le \alpha = 1/2, \\ \sigma_{n+1}\sqrt{\ln(n+1)} & \text{if } \beta > \alpha = 1/2, \\ \sigma_{n+1} & \text{if } \alpha > 1/2. \end{cases}$$

Diameters of sections – previous results

- diameters of sections of symmetric convex bodies with a lower-dimensional subspace have been studied in asymptotic geometric analysis
- initiated by Giannopoulos and Milman (1997,1998)
- subsequent work of Litvak and Tomczak-Jaegermann (2000), Giannopoulos, Milman, and Tsolomitis (2005), Litvak, Pajor, and Tomczak-Jaegermann (2006)

Diameters of sections – previous results

- these bounds are not sharp for the whole class of symmetric convex bodies, in particular not for ellipsoids with highly incomparable semi-axes
- focus in these papers was on subspaces of proportional codimension, whereas we are mainly interested in subspaces with small codimension such as $m = n^2$ or $m = 2^n$
- incidentally, one of our first approaches based on entropy numbers used ideas from Litvak, Pajor, and Tomczak-Jaegermann (2006), but did not give sharp results

Upper bound

Theorem

Let $\sigma \in \ell_2$ be non-increasing. Then, for all $n \in \mathbb{N}$, we have

$$\mathbb{P}\left[r(S, N_n) \geq \frac{221}{\sqrt{n}} \left(\sum_{j \geq \lfloor n/4 \rfloor} \sigma_j^2 \right)^{1/2} \right] \leq 2e^{-n/100}$$

Upper bound - tools

- Let $G_n = (g_{ij})_{1 \leq i \leq n, j \in \mathbb{N}}$ have independent standard Gaussian entries.
- We want to study the distribution of the random variable

$$r(S, N_n) = \sup \left\{ \|x\|_2 : x \in \mathcal{E}_\sigma,\ G_n(x) = 0 \right\}.$$

- For index sets $I, J \subseteq \mathbb{N}$, we consider the (structured) Gaussian $I \times J$-matrices

$$G_{I,J} = (g_{ij})_{i \in I, j \in J} \quad \text{and} \quad \Sigma_{I,J} = (\sigma_j g_{ij})_{i \in I, j \in J}.$$

Upper bound - tools - Radius vs singular numbers

Proposition

Let $\sigma \in \ell_2$ be non-increasing and let $k \leq n$. If $G_{n,k} \in \mathbb{R}^{n\times k}$ has full rank, then

$$r(S, N_n) \leq \sigma_{k+1} + \frac{s_1\left(\Sigma_{[n],\mathbb{N}\setminus[k]}\right)}{s_k\left(G_{n,k}\right)}.$$

- then bound k-th, i.e., smallest, singular value of the Gaussian matrix $G_{n,k}$ from below (using Bandeira and Van Handel 2016) and ...
- ... and the largest singular value of the structured Gaussian matrix $\Sigma_{[n],\mathbb{N}\setminus[k]}$ from above (using Davidson and Szarek (2001))

Lower bound

Proposition

Let $\sigma \in \ell_2$ be non-increasing, $\varepsilon \in (0,1)$, and $n, k \in \mathbb{N}$ be such that $\sigma_k \neq 0$ and

$$\sum_{j>k} \sigma_j^2 \geq \frac{3n\sigma_k^2}{\varepsilon^2}.$$

Then

$$\mathbb{P}\Big[r(S, N_n) \ \leq \ \sigma_k(1-\varepsilon)\Big] \leq 5e^{-n/64}.$$

Lower bound - tools

- again we prove a lower bound including singular values and norms of random matrices
- then Gordon's min-max theorem (Gordon (1988)) and concentration of chi-square random variables (Laurent and Massart (2000)) or concentration of Gaussian random vectors in Banach spaces can be used

Analogy

Results are similar for random information and for information based on function evaluations. Also for the latter case it is very important whether $\sigma \in \ell_2$ or not.

See results of Wasilkowski, Woźniakowski (2001),
Kuo, Wasilkowski and Woźniakowski (2009)
and Hinrichs, Novak and Vybíral (2008).
See the book of Novak and Woźniakowski (Part 3, 2012).