How good is random information?

joint work with Aicke Hinrichs, David Krieg, Joscha Prochno and Mario Ullrich

Erich Novak

FSU Jena

Linz, November 2018

Problems and Algorithms - IBC Setting

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

Error and Radius of Information

- **Worst Case Error:** $e(A_n) = \sup\{\|Sf - \varphi(N_n f)\|_G : \|f\|_F \leq 1\}$
- **Radius of Information:** $r(N_n) = \inf_{\varphi} e(\varphi \circ N_n)$
- **Radius at Zero:** $r_0(N_n) = \sup\{\|Sf\|_G : \|f\|_F \leq 1, N_n f = 0\}$
- **Result:** $r_0(N_n) \leq r(N_n) \leq 2r_0(N_n)$
- **Minimal Worst Case Error:** $e_n = \inf_{\varphi, N_n} e(\varphi \circ N_n) = \inf_{N_n} r(N_n)$
- **Optimal Information:** N_n^* with $r(N_n^*) = e_n$

Random Information vs. Optimal Information

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

Results for Function Evaluations, Λ^{std}

Here we consider functions $f\colon [0,1]^d \to \mathbb{R}$ and take i.i.d. info of the form $L_i(f) = f(x_i)$ with uniformly distributed $x_i \in [0,1]^d$.

Depending on F and S another distribution of the x_i might be better, "importance sampling".

In the randomized setting studied by Plaskota, Wasilkowski and Hinrichs. See our book, Section 17.2.

1. Example: Lipschitz functions, $d = 1$

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

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

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

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

2. Example: Discrepancy for $d = 1$

Consider p-discrepancy or integration for F_q,

$F_q = \{f : [0,1] \to \mathbb{R} \; : \; f(0) = 0, \; \|f'\|_q \leq 1\}, \;\; 1/p + 1/q = 1.$

Again, the numbers e_n are between $\frac{1}{2n}$ (for $q = 1$, star discrepancy) and $\frac{1}{4n}$ (for $q = 1$).

Heidi Weyhausen proved in her dissertation (2015):

For the star discrepancy ($q = 1$):

$$c \cdot \frac{\log n}{n} \leq \mathbb{E}(r(N_n)) \leq \frac{8 \log n}{n},$$

and for $p = q = 2$:

$$\mathbb{E}(r(N_n)^2) \asymp n^{-2}.$$

3. Example: Imbedding of Sobolev spaces

$S = Id \colon F \to G$ with $G = L_q$ and $F = W_p^k([0,1]^d)$, where $kp > d$.

The order $e_n \asymp n^{-\alpha}$ is known: $\alpha = k/d - (1/p - 1/q)_+$.

We *know* that in the case $q = \infty$

$$\mathbb{E}(r(N_n)) \asymp \left(\frac{\log n}{n}\right)^{\alpha}$$

and we *conjecture* that

$$e_n \asymp \mathbb{E}(r(N_n))$$

for $q < \infty$ and $p > 1$ (we have a proof for Lipschitz functions, $k = 1$ and $p = \infty$).

Results for Linear Functionals, $\Lambda^{\rm all}$

We consider two cases:

1) $S = Id \colon \ell_1^m \to \ell_2^m$, we present well known results.

2) $S = Id \colon F \to L_2$, where F is a Hilbert space, compactly embedded into a separable L_2-space. Here we have new results.

Result of Kashin (1975) and Gluskin (1983)

For the problem of ℓ_2^m approximation of vectors in ℓ_1^m, random information is almost optimal.

Random means: $L_i x = \langle x, y_i \rangle$, where $y_1, \ldots, y_n$ are i.i.d. uniform on the Euclidean unit sphere of $\mathbb{R}^m$.

Let $S = Id \colon \ell_1^m \to \ell_2^m$. Then

$$e_n = r(N_n^*) \asymp \mathbb{E}(r(N_n)) \asymp \sqrt{\frac{\log(m/n)}{n}},$$

for $m > 2n$.

L_2-approx of functions from a Hilbert space

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

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

Optimal Information ...

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

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

Random Information

- We want to compare this to random Gaussian information given by $L_i x = \sum_{j=1}^{m} g_{ij} x_j$ ...
- ... where g_{ij} are i.i.d. standard Gaussian random variables
- equivalently for finite m: $L_i x = \langle x, y_i \rangle$ 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 $m = \infty$: $L_i x = \sum_{j=1}^{m} g_{ij} x_j$ can be undefined.
- But: If $\sigma \in \ell_2$ then $L_i \colon F \to \mathbb{R}$ is finite almost surely for $i = 1, \dots, n$ and

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

  Better bound:

$$\mathbb{E} \| N_n \colon F \to \ell_2^n \| \le C \cdot \max\{\sqrt{n}, \|\sigma\|_2\}.$$

- 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 $\mathbb{R}^n$.

Theorem (Case $m = \infty$)

$$\sigma \in \ell_2 \implies \mathbb{E} r(N_n) \prec n^{-1/2}$$

$$\sigma \notin \ell_2 \implies \mathbb{E} r(N_n) = 1 \quad \textit{for } n \in \mathbb{N}$$

The main upper bound for $\sigma \in \ell_2$ is:

$$r(N_n) \leq 8\, \sigma_{n/4} + 7\, \frac{\sqrt{\sum_{j=n/4}^{\infty} \sigma_j^2}}{\sqrt{n}}$$

with probability at least $1 - 2\, e^{-n/32}$.

Consequence

- now $m \in \mathbb{N} \cup \{\infty\}$
- exemplary case: $\sigma_n = n^{-\alpha}$

Theorem

$$\alpha > \frac{1}{2} \implies \mathbb{E}\, r(N_n) \asymp \sigma_n = n^{-\alpha}$$
$$\alpha < \frac{1}{2} \implies r(N_n) \succ 1 \textit{ for } n \leq m^{1-2\alpha} \textit{ with high probability}$$

Lower bound

$$r(N_n) \geq 1 - c\,\frac{\sqrt{n}}{\|\sigma\|_2} \qquad \text{with large probability}$$

In particular:

- Random info is bad for $n \prec \|\sigma\|_2^2$
- Random info is bad for $n \prec \log m$ if $\sigma_n = n^{-1/2}$
- $r(N_n) = 1$ a.s. if $\sigma \notin \ell_2$.

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, Chapter 26, 2012).

- exemplary case: $\sigma_n = n^{-\alpha}$ where $\alpha > 1/2$

Kuo, Wasilkowski, Woźniakowski (2009) conjecture that the order $n^{-\alpha}$ can also be achieved with Λ^{std}.

They proved that the following order can be achieved:

$$n^{-\frac{2\alpha^2}{2\alpha+1}}.$$