Randomness in functional analysis: toward universality

Roman Vershynin

The probabilistic method

- Allows us to find objects that are hard to construct deterministically.
- Suggests typical objects.
- Combinatorics (Erdös): random graphs.
- Functional analysis: random spaces and operators.

Banach versus Hilbert spaces

- Hilbert spaces: simple geometry.
  Banach spaces: complicated geometry.
- Can we reduce a Banach space to a Hilbert space?

Does every infinite-dimensional Banach space contain a infinite-dimensional Hilbertian subspace?

- No. ℓ_p, $1 \leq p < \infty$.

A Hilbertian subspace of *arbitrary large* dimension?

- Yes. Dvoretzky's theorem (1960) answering a question by Grothendieck.

Hilbertian subspaces

Dvoretzky's Theorem [Dvoretzky '60, Milman '71]

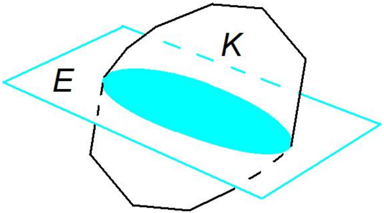
Every Banach space X of dimension n contains a Hilbertian subspace E of dimension $C(n) \to \infty$ as $n \to \infty$. Random subspace works.

- Correspondence with convex geometry:
- $K =$ the unit ball of X. Arbitrary symmetric convex body.

Spherical sections of convex bodies

Dvoretzky's Theorem (geometric form)

Every symmetric convex body K in $\mathbb{R}^n$ has an almost round section $K \cap E$ of dimension $C(n) \to \infty$ as $n \to \infty$. Random section works.

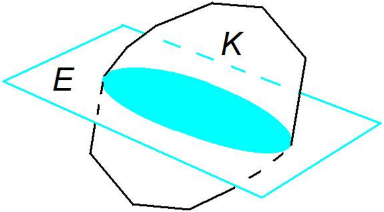


- The probabilistic method must work if any other method works:

The probabilistic method must always work

Randomness Theorem [V '04], [Giannopoulos, Milman, Tsolomitis '04]

Assume that some $(n/2)$-dimensional section of K has diameter ≤ 1.
Then a random $(n/4)$-dimensional section of K has diameter $\leq C$.



- Both Dvoretzky's Theorem and Randomness Theorems rely on isoperimetric inequalities for the sphere: Lévy's and Gromov's resectively.

How large are Hilbertian subspaces?

What is the dimension $c(n)$ of the largest Hilbertian subspace of X?

- Worst space:

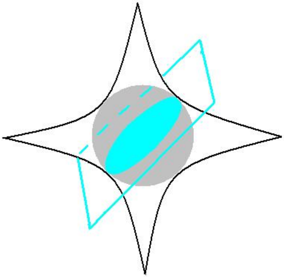
$$X = \ell_\infty \qquad \|x\|_\infty = \max_k |x_k|$$

  Then $c(n) \sim \log n$. For any other space, $c(n)$ is bigger [Milman '71].
- Best space:

$$X = \ell_1 \qquad \|x\|_1 = |x_1| + \cdots + |x_n|$$

  Then $c(n) \sim n$, even $0.99n$. [Kashin '77].

- Heuristics: Ball(X) $=$ round $+$ few spikes.
  Random subspace avoids the spikes,
  only feels the round part.
- General spaces X: true
  (provided volume is mostly in the round part)



Universality

- Most constructions in geometric functional analysis are random.
- Does the specific distribution play any role? We hope not.
- Universality: macroscopic behavior of a random system is independent of the distributions of its microscopic parts.
- Basic illustration: Central Limit Theorem.

Universality: Hilbertian subspaces

Universality Conjecture

For any distribution on the subspaces (sufficiently rich), random subspaces of Banach spaces are Hilbertian.

- Natural representation of random subspaces?
  As images of random operators.
  Random matrices with random independent entries.
- Universality Conjecture: random operators embed Hilbert spaces into Banach spaces. True for $X = \ell_1$ (and the like):

Universality Theorem [Litvak, Pajor, Rudelson, Tomczak, V '05]

With high probability, a random operator $A : \mathbb{R}^{0.99n} \to \mathbb{R}^n$ satisfies

$$cn \le \|Ax\|_1 \le Cn \qquad \text{for all } x,\ \|x\|_2 = 1.$$

- This proves a conjecture of Schechtman.

Universality: random operators

- Let's try to understand random operators in the most basic case: on finite dimensional Hilbert spaces.
- Basic questions: boundedness, invertibility.

Boundedness/Invertibility Problem

Consider a random operator $A : \mathbb{R}^n \to \mathbb{R}^n$, whose matrix has independent entries. What are the norms of A and A^{-1}?

- Boundedness: $\|A\| \sim \sqrt{n}$ with high probability [Von Neumann, Goldstine '47].
- Universality: the limiting distribution of $\|A\|$ (properly normalized) is independent of the distribution of the entries [Soshnikov '02].
- Invertibility: much harder. It is even nontrivial that A is invertible with high probability [Komlos '67]. Universality is unknown.

Universality: random operators

Conjecture [Von Neumann '47], [Smale '85], [Spielman-Teng '02]

Random operators satisfy $\|A^{-1}\| \sim \sqrt{n}$ with high probability.

- In other words, random operators are good isomorphisms:

$$\frac{1}{\sqrt{n}} \lesssim \|Ax\|_2 \lesssim \sqrt{n} \quad \text{for all unit vectors } x.$$

- Smale's Conjecture: true for Gaussian A.
  Proved by [Edelman '88], [Szarek '90].
- Spielman-Teng's Conjecture (ICM 2002): true for Bernoulli A.

Universality: random operators

Invertibility Theorem [Rudelson-V '07]

Von Neumann's Conjecture is true. For a random operator A on $\mathbb{R}^n$ (matrix with independent entries), $\|A^{-1}\| \lesssim \sqrt{n}$ with high probability.

- Partial results: [Komlos '67], [Edelman '88], [Szarek '90], [Kahn-Komlos-Szemeredi '95], [Rudelson '06], [Tao-Vu '06].
- Theorem is a first step toward universality for $\|A^{-1}\|$.
- A similar result for random embeddings:

Universality: random operators

Invertibility Theorem [Rudelson-V '08]

For a random operator $A : \mathbb{R}^n \to \mathbb{R}^N$ (matrix with independent entries),

$$\sqrt{N} - \sqrt{n} \lesssim \|Ax\|_2 \lesssim \sqrt{N} + \sqrt{n} \quad \text{for all unit vectors } x.$$

- Previous results: true in the limit as $n \to \infty$, $N/n \to$ const [Silverstein '85] (for Gaussian), [Bai-Yin '93].
- Partial results: [Litvak, Pajor, Rudelson, Tomczak '05], [Artstein, Friedland, Milman, Sodin '06] (analyzing Bai-Yin), [Rudelson '06].

Techniques and Connections

Random operators

↓

Geometry (distance problems)

↓

Probability Theory (sums of independent random variables)

↓

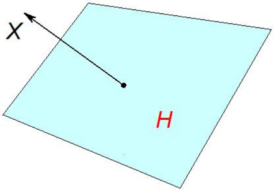
Additive Combinatorics (arithmetic progressions)

↓

Ergodic Theory (recurrence)

Geometry: the distance problem

- How to show that an operator A on $\mathbb{R}^n$ is well invertible?
- A is not invertible $\iff$ one of its columns X lies in the span H of the others.



- So, the invertibility problem reduces to a *geometric* problem:

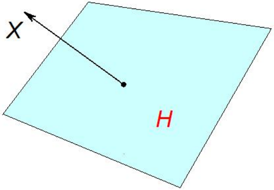
Distance Problem

What is the distance in $\mathbb{R}^n$ between a random vector X and a random hyperplane H?

Geometry: the distance problem

Distance Problem

What is the distance in $\mathbb{R}^n$ between a random vector X and a random hyperplane H?



- Easy case: Gaussian distribution. The answer is constant:

$$\text{dist}(X, H) = |\text{Gaussian}| \sim \text{const}$$

  with high probability. Precisely, for every $\varepsilon > 0$,

$$\mathbb{P}(\text{dist}(X, H) \le \varepsilon) \sim \varepsilon.$$

- Is this true in general? Yes. Non-trivial even for Bernoulli:

Geometry: the distance problem

Distance Theorem

The distance between a random vector X and a random hyperplane H in $\mathbb{R}^n$ is $\sim$ constant with high probability. Precisely, for $\varepsilon > c^n$,

$$\mathbb{P}(\mathrm{dist}(X, H) \le \varepsilon) \lesssim \varepsilon.$$

- Partial result: [Tao-Vu '06].
- Applied to the columns of a random matrix, Distance Theorem immediately yields:

Invertibility Theorem

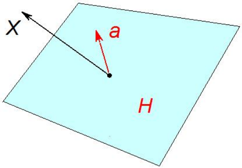
A random operator A on $\mathbb{R}^n$ is invertible with high probability $1 - c^n$.

- Partial result: [Kahn-Komlos-Szemeredi '95] for Bernoulli.

Probability Theory: sums of random variables

- Distance Problem → probability theory:
  represent $\text{dist}(X, H)$ as a sum of independent random variables.
- Condition on the hyperplane H with normal a. Then

$$\text{dist}(X, H) = |\langle a, X \rangle| = \left| \sum_{k=1}^{n} a_k \xi_k \right|$$



where $a = (a_1, \ldots, a_n), \quad X = (\xi_1, \ldots, \xi_n)$.

- Then the Distance Problem translates to a problem about the *sum of independent random variables*

$$S = \sum_{k=1}^{n} a_k \xi_k .$$

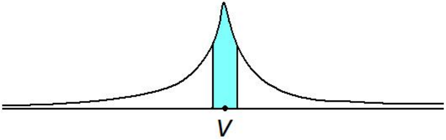
Want to bound S *below*.

Probability Theory: concentration

Definition [P. LÉVY]

The concentration function of a random variable S is

$$\mathrm{conc}(\varepsilon) = \sup_v \mathbb{P}(|S - v| \le \varepsilon).$$

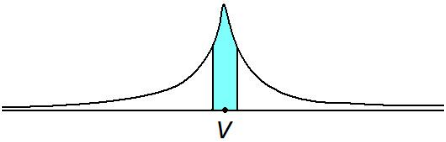


- The concentration function measures how spread S is.
- Ideal case: S has bounded density; then

$$\mathrm{conc}(\varepsilon) \sim \varepsilon \qquad (*)$$

- Opposite to large deviation, concentration inequalities.

Probability Theory: concentration



$$\text{conc}(\varepsilon) = \sup_{v} \mathbb{P}(|S - v| \leq \varepsilon).$$

Concentration Problem

Estimate the concentration function of a *sum of independent random variables* $S = \sum a_k \xi_k$. When is $\text{conc}(\varepsilon) \sim \varepsilon$?

- Early results: Lévy, Kolmogorov, Esséen, Littlewood-Offord, Erdös.
- Soft approach: use Central Limit Theorem to approximate S by a Gaussian random variable.
- But the rate of convergence in CLT is too slow [Berry-Esséen].
- Other approaches suggested in connection with additive combinatorics, analytic number theory.

Additive Combinatorics: Littlewood-Offord Problem

Littlewood-Offord Problem

Given nonzero real numbers $a_1, \ldots, a_n$, how many sign-sums

$$S = \sum \pm a_k$$

can be equal zero?

- This is a partial case of the Concentration Problem. View $\pm$ as random signs. Hence at most $\mathrm{conc}(0) \cdot 2^n$.

Additive Combinatorics: Littlewood-Offord Problem

- Worst case for $S = \sum \pm a_k$ is $(a_k) = (1, 1, \ldots, 1)$:

Theorem [Littlewood-Offord '43], [Erdös '45]

If all $|a_k| \geq 1$ then $\text{conc}(1) \lesssim 1/\sqrt{n}$ (middle binomial).

- This is too big! Remove the structure, make a_k incomparable:

Theorem [Erdös-Moser '65], [Sárkozi-Szemerédi '65], [Hálasz '77]

If all $|a_j - a_k| \geq 1$ then $\text{conc}(1) \lesssim 1/n^{3/2}$.

- Sharp for $(a_k) = (1, 2, 3, \ldots, n)$. Still big! Lot of structure.
- Remove structure... $\text{conc}(0) \lesssim$ arbitrary polynomial [Tao-Vu '06].
- New result: can reduce to arbitrary function:

Additive Combinatorics: Littlewood-Offord Problem

Structure Theorem [Rudelson-V. '07]

A sum of independent random variables $S = \sum a_k \xi_k$ with $\|a\|_2 = 1$ satisfies one of the following:

Universality: $\text{conc}(\varepsilon) \lesssim \varepsilon$

Structure: the coefficients (a_k) can be essentially embedded into a short arithmetic progression, length $\sim 1/\,\text{conc}(\varepsilon)$.

- In other words, the obstacle in the Littlewood-Offord Problem is additive structure of the coefficients (arithmetic progressions).
- By removing structure, we achieve the universality.
- Structure Theorem covers (asymptotically) the results of [Littlewood-Offord '43], [Erdös '45], [Erdös-Moser '65], [Sárkozi-Szemerédi '65], [Hálasz '77].
- What coefficient sequences (a_k) have no structure? For example, random coefficients.
- Example: $a = (a_k)$ is the normal of a random hyperplane in $\mathbb{R}^n$. (Recall our invertibility problem for random operators.)

Additive Combinatorics: Littlewood-Offord Problem

No Structure Theorem [Rudelson-V. '07]

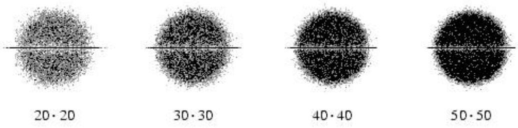
The normal a of a random hyperplane H in $\mathbb{R}^n$ (spanned by independent vectors) can not be essentially embedded into an arithmetic progression of length smaller than C^n.

- Nontrivial: the coefficients of a are not independent.
- Going backward, one deduces the Invertibility Theorem for random operators.

Application of the Invertibility Theorem: Circular Law

Circular Law

Let A be an $n \times n$ random matrix with real independent entries. Then the joint distribution of the eigenvalues of $\frac{1}{\sqrt{n}}A$ converges as $n \to \infty$ to the uniform distribution on the unit disc in $\mathbb{C}$.



History:

- Suggested by [PASTUR '73]. Claimed by [GIRKO '84].
- [EDELMAN '97] for Gaussian matrices
- [BAI '97] assuming density, 6-th moments of the entries
- [GÖTZE-TIKHOMIROV '07], [TAO-VU '07]: general matrices with $2 + o(1)$ moments. Deduced from the Invertibility Theorem.

Techniques and Connections

Random operators

↓

Geometry (distance problems)

↓

Probability Theory (sums of independent random variables)

↓

Additive Combinatorics (arithmetic progressions)

↓

Ergodic Theory (recurrence)