On the geometry of random polytopes generated by heavy-tailed random vectors

Holger Rauhut
Chair for Mathematics of Information Processing
RWTH Aachen University

Asymtotic Geometric Analysis IV
Euler Institute Sankt Petersburg
July 2, 2019

Joint work with
O. Guédon, F. Krahmer, C. Kümmerle, S. Mendelson

Centrally symmetric random polytopes

X: symmetric random vector in $\mathbb{R}^n$

Given independent copies $X_1, \ldots, X_N$ of X define random polytope

$$P_N(X) := \text{absconv}(X_1, \ldots, X_N) = \text{conv}(\pm X_1, \ldots, \pm X_N)$$
$$= \left\{ \sum_{j=1}^{N} \alpha_j X_j : \alpha_j \in [-1, 1], \sum_{j=1}^{N} |\alpha_j| \leq 1 \right\}$$

Can we find a (deterministic) large "canonical body" $K \subset \mathbb{R}^n$ s.t.

$$K \subset P_N(X) \quad \text{with high probability}$$

under very general and weak conditions on X?

Description of K as intersection of ℓ_p-balls?

Two notable results

Theorem (Gluskin 1989)

Let X *be a standard* *Gaussian* *random vector in* $\mathbb{R}^n$*, set* $0 < \alpha < 1$ *and consider* $N \geq c_0(\alpha)n$*. Then*

$$c_1(\alpha)\sqrt{\log(eN/n)}B_2^n \subset \mathrm{absconv}(X_1, \ldots, X_N)$$

with probability at least $1 - 2\exp(-c_2 N^{1-\alpha} n^{\alpha})$*.*

Two notable results

Theorem (Gluskin 1989)

Let X be a standard Gaussian random vector in $\mathbb{R}^n$, set $0 < \alpha < 1$ and consider $N \geq c_0(\alpha)n$. Then

$$c_1(\alpha)\sqrt{\log(eN/n)}B_2^n \subset \operatorname{absconv}(X_1, \dots, X_N)$$

with probability at least $1 - 2\exp(-c_2 N^{1-\alpha} n^{\alpha})$.

Theorem (Giannopoulos, Hartzoulaki 2002; Litvak, Pajor, Rudelson, Tomczak-Jaegermann 2005)

Let ξ be a mean-zero, unit variance, L-subgaussian random variable and set $X = (\xi_i)_{i=1}^n$. For $0 < \alpha < 1$, consider $N \geq c_0(\alpha, L)n$. Then with probability at least $1 - 2\exp(-c_1 N^{1-\alpha} n^{\alpha})$

$$c_2(\alpha, L)\big(B_\infty^n \cap \sqrt{\log(eN/n)}B_2^n\big) \subset \operatorname{absconv}(X_1, \dots, X_N).$$

In the case of Rademacher vector $X = \mathcal{E}$, the theorem is false without the intersection with the unit ball in ℓ_∞.

Floating Bodies

Floating body associated to symmetric random vector X in $\mathbb{R}^n$:

$$K_p(X) := \{t \in \mathbb{R}^n : \mathbb{P}(\langle X, t\rangle \geq 1) \leq \exp(-p)\}, \quad p \geq 1.$$

Floating Bodies

Floating body associated to symmetric random vector X in $\mathbb{R}^n$:

$$K_p(X) := \{t \in \mathbb{R}^n : \mathbb{P}(\langle X, t\rangle \geq 1) \leq \exp(-p)\}, \quad p \geq 1.$$

Polar body of $T \subset \mathbb{R}^n$:

$$T^\circ = \{x \in \mathbb{R}^n : \langle t, x\rangle \leq 1 \ \text{ for every } t \in T\}$$

Floating Bodies

Floating body associated to symmetric random vector X in $\mathbb{R}^n$:

$$K_p(X) := \{t \in \mathbb{R}^n : \mathbb{P}(\langle X, t\rangle \geq 1) \leq \exp(-p)\}\,, \quad p \geq 1.$$

Polar body of $T \subset \mathbb{R}^n$:

$$T^\circ = \{x \in \mathbb{R}^n : \langle t, x\rangle \leq 1 \text{ for every } t \in T\}$$

- If X is standard Gaussian in $\mathbb{R}^n$ then

$$K_p(X) \sim \frac{1}{\sqrt{p}} B_2^n \quad \text{and} \quad (K_p(X))^\circ \sim \sqrt{p} B_2^n$$

- If $X = \mathcal{E}$ is standard Rademacher in $\mathbb{R}^n$ then

$$K_p(\mathcal{E}) \sim \mathrm{conv}(B_1^n \cup (1/\sqrt{p}) B_2^n) \quad \text{and} \quad (K_p(\mathcal{E}))^\circ \sim B_\infty^n \cap \sqrt{p} B_2^n.$$

Floating Bodies

Floating body associated to symmetric random vector X in $\mathbb{R}^n$:

$$K_p(X) := \{t \in \mathbb{R}^n : \mathbb{P}(\langle X, t\rangle \geq 1) \leq \exp(-p)\}, \quad p \geq 1.$$

Polar body of $T \subset \mathbb{R}^n$:

$$T^\circ = \{x \in \mathbb{R}^n : \langle t, x\rangle \leq 1 \text{ for every } t \in T\}$$

- If X is standard Gaussian in $\mathbb{R}^n$ then

$$K_p(X) \sim \frac{1}{\sqrt{p}} B_2^n \quad \text{and} \quad (K_p(X))^\circ \sim \sqrt{p} B_2^n$$

- If $X = \mathcal{E}$ is standard Rademacher in $\mathbb{R}^n$ then

$$K_p(\mathcal{E}) \sim \operatorname{conv}(B_1^n \cup (1/\sqrt{p}) B_2^n) \quad \text{and} \quad (K_p(\mathcal{E}))^\circ \sim B_\infty^n \cap \sqrt{p} B_2^n.$$

In both cases, for $p = \alpha \log(eN/n)$, with high probability

$$c(K_p(X))^\circ \subset \operatorname{absconv}(X_1, \ldots, X_N).$$

Does this inclusion extend to more general random vectors?

Assumptions on X

For some norm $\|\cdot\|$ on $\mathbb{R}^n$, assume that the symmetric random vector X satisfies

- the small ball condition

$$\mathbb{P}(|\langle X, t\rangle| \geq \gamma\|t\|) \geq \delta \quad \text{for all } t \in \mathbb{R}^n$$

  for some constants $\gamma, \delta > 0$.

- the L_r condition

$$(\mathbb{E}|\langle X, t\rangle|^r)^{1/r} \leq L\|t\| \quad \text{for all } t \in \mathbb{R}^n$$

  for some $r > 0$ and some constant $L > 0$.

Assumptions on X

For some norm $\|\cdot\|$ on $\mathbb{R}^n$, assume that the symmetric random vector X satisfies

- the small ball condition

$$\mathbb{P}(|\langle X, t\rangle| \geq \gamma\|t\|) \geq \delta \quad \text{for all } t \in \mathbb{R}^n$$

  for some constants $\gamma, \delta > 0$.

- the L_r condition

$$(\mathbb{E}|\langle X, t\rangle|^r)^{1/r} \leq L\|t\| \quad \text{for all } t \in \mathbb{R}^n$$

  for some $r > 0$ and some constant $L > 0$.

Note that assumption

- allows very heavy-tailed random vectors
- does not require independence of the entries of X
- does not require isotropic random vectors X

Main Theorem

Theorem (Guédon, Krahmer, Kümmerle, Mendelson, R 2019)

Let X be a symmetric random vector that satisfies the assumption with respect to a norm $\|\cdot\|$ and some $\delta, \gamma, r > 0$. Let $0 < \alpha < 1$ and set $p = \alpha \log(eN/n)$. If

$$N \geq c_0 n \quad \text{for } c_0 = c_0(\alpha, \gamma, \delta, r, L),$$

then with probability at least $1 - 2\exp(-c_1 N^{1-\alpha} n^{\alpha})$,

$$c_2 \big(K_p(X)\big)^{\circ} \subset \mathrm{absconv}(X_1, \ldots, X_N),$$

where c_1 and c_2 are absolute constants.

Main Theorem

Theorem (Guédon, Krahmer, Kümmerle, Mendelson, R 2019)

Let X be a symmetric random vector that satisfies the assumption with respect to a norm $\|\cdot\|$ and some $\delta, \gamma, r > 0$. Let $0 < \alpha < 1$ and set $p = \alpha \log(eN/n)$. If

$$N \geq c_0 n \qquad \text{for } c_0 = c_0(\alpha, \gamma, \delta, r, L),$$

then with probability at least $1 - 2\exp(-c_1 N^{1-\alpha} n^{\alpha})$,

$$c_2 \big(K_p(X)\big)^{\circ} \subset \mathrm{absconv}(X_1, \ldots, X_N),$$

where c_1 and c_2 are absolute constants.

Recovers or improves previous results

- X Gaussian or subgaussian (as discussed)
- X isotropic log-concave (Dafnis, Giannopoulos, Tsolomitis 2009)
- X with independent coordinates satisfying a small ball condition (Krahmer, Kümmerle, R 2018; Guédon, Litvak, Tatarko 2018; Mendelson 2019)

Example: q-stable random vectors

X: q-stable random vector, i.e., with independent q-stable entries $(1 \leq q < 2)$

For $1 \leq q < 2$ a standard q-stable random variable ξ is defined via its characteristic function

$$\mathbb{E}[\exp(it\xi)] = \exp(-|t|^q/2) \quad \text{for all } t \in \mathbb{R}.$$

Example: q-stable random vectors

X: q-stable random vector, i.e., with independent q-stable entries $(1 \leq q < 2)$

For $1 \leq q < 2$ a standard q-stable random variable ξ is defined via its characteristic function

$$\mathbb{E}[\exp(it\xi)] = \exp(-|t|^q/2) \quad \text{for all } t \in \mathbb{R}.$$

Recall:

- $\sup_{u>0} u^q \mathbb{P}(|\xi| > u) \leq C_q$, and $\mathbb{P}(|\xi| > u) \geq c_q/u^q$ for $u \geq M_q$.
- If $\xi_1, \ldots, \xi_n$ are independent copies of ξ then $\sum_{i=1}^n t_i\xi_i$ has the same distribution as $\|t\|_q \, \xi$ for all $t \in \mathbb{R}^n$ (stability property)

$q = 1$: Cauchy variable (expectation does not exist)

Theorem for q-stable random vectors

Theorem

Let X be a q-stable random vector on $\mathbb{R}^n$ for some $1 \le q < 2$ and let $X_1, \ldots, X_N$ be independent copies of X. Then for $0 < \alpha < 1$ and $N \ge c_0(\alpha, q)n$, with probability at least $1 - 2\exp(-c_1 N^{1-\alpha} n^{\alpha})$,

$$c_2(\alpha, q) \left(\frac{N}{n}\right)^{\alpha/q} B^n_{q'} \subset \mathrm{absconv}(X_1, \ldots, X_N),$$

where $1/q + 1/q' = 1$.
In particular, if X is a Cauchy random vector ($q = 1$), then with probability at least $1 - 2\exp(-c_1 N^{1-\alpha} n^{\alpha})$

$$c_3(\alpha) \left(\frac{N}{n}\right)^{\alpha} B^n_{\infty} \subset \mathrm{absconv}(X_1, \ldots, X_N).$$

$K_p(X)$ as L_p-centroid body

For X with sufficiently many moments

$$B(L_p(X)) := \{t \in \mathbb{R}^n : (\mathbb{E}|\langle X, t\rangle|^p)^{1/p} \leq 1\}$$

L_p-centroid body

$$Z_p(X) := B_p(X)^\circ$$

$K_p(X)$ as L_p-centroid body

For X with sufficiently many moments

$$B(L_p(X)) := \{t \in \mathbb{R}^n : (\mathbb{E}|\langle X, t\rangle|^p)^{1/p} \leq 1\}$$

L_p-centroid body

$$Z_p(X) := B_p(X)^\circ$$

If X satisfies the regularity condition

$$\|\langle t, X\rangle\|_{L_{2q}} \leq L\|\langle t, X\rangle\|_{L_q} \quad \text{for all } t \in \mathbb{R}^n \text{ and all } q \geq 2$$

then (first part consequence of Markov's inequality)

$$e^{-1}B(L_p(X)) \subset K_p(X) \subset 2B(L_{cp}(X)).$$

Log-concave random vectors

Theorem

Let X be a symmetric logarithmically concave random vector. Let $0 < \alpha < 1$, set $N \geq c_0(\alpha)n$ and put $p = \alpha \log(eN/n)$. Then, with probability at least $1 - 2\exp(-c_1 N^{1-\alpha} n^{\alpha})$,

$$\mathrm{absconv}(X_1, \ldots, X_N) \supset c_2(\alpha) Z_p(X).$$

Improvement over probability estimate due to Dafnis, Giannopolous, Tsolomitis (2009):

$$1 - 2\exp(-c_1 N^{1-\alpha} n^{\alpha}) - \mathbb{P}\left(\|\Gamma : \ell_2^n \to \ell_2^N\| \geq c\sqrt{N}\right), \qquad (1)$$

where $\Gamma = (X_1 | \cdots | X_N)^T$. (In case of independent entries (1) is implied by Litvak, Pajor, Rudelson, Tomczak-Jaegermann 2005.)

Independent entries satisfying small ball assumption

Theorem (Guédon, Litvak, Tatarko 2018)

Let x be a symmetric random variable satisfying $\mathbb{E}x^2 = 1$, let $x_1, \ldots, x_n$ to be independent copies of x and put $X = (x_i)_{i=1}^n$. If there are constants $\gamma, \delta > 0$ such that $\mathbb{P}(|x| \geq \gamma) \geq \delta$, then for $N \geq c_0 n$, with probability at least $1 - 2\exp(-c_1 N^{1-\alpha} n^{\alpha})$,

$$\mathrm{absconv}(X_1, \ldots, X_N) \supset c_2(B_\infty^n \cap \sqrt{\log(eN/n)} B_2^n).$$

Here c_0 and c_2 depend on α, γ and δ, and c_1 is an absolute constant.

Unconditional random vectors

A random vector $X = (x_i)_{i=1}^n$ is unconditional if for every vector $(\varepsilon_i)_{i=1}^n \in \{-1, 1\}^n$, $(x_i)_{i=1}^n$ has the same distribution as $(\varepsilon_i x_i)_{i=1}^n$.

Theorem

Let X be an unconditional random vector that satisfies the small-ball condition with constants γ and δ. Then, for any $p > c_0(\delta) = 4\log(8/\delta) + \log(4)$,

$$K_p(X) \subset \frac{c(\delta)}{\gamma} K_p(\mathcal{E}) \sim \frac{c(\delta)}{\gamma}(B_\infty^n \cap \sqrt{\log(eN/n)} B_2^n).$$

In particular, if X satisfies our assumption and $N \geq c_0(\alpha, \gamma, \delta, r, L)n$, *then with probability at least* $1 - 2\exp(-c_1 N^{1-\alpha} n^\alpha)$,

$$\mathrm{absconv}(X_1, \ldots, X_N) \supset c'(\delta)\gamma\big(B_\infty^n \cap \sqrt{\log(eN/n)} B_2^n\big).$$

Implication for compressive sensing

Compressive sensing problem:
Recover an (approximately) s-sparse vector $x \in \mathbb{R}^N$, i.e.,

$$\|x\|_0 = \#\{\ell : x_\ell \neq 0\} \leq s$$

from n linear noisy measurements

$$y = Ax + w \quad \text{with } A \in \mathbb{R}^{n \times N} \quad \text{where } n \ll N.$$

Implication for compressive sensing

Compressive sensing problem:
Recover an (approximately) s-sparse vector $x \in \mathbb{R}^N$, i.e.,

$$\|x\|_0 = \#\{\ell : x_\ell \neq 0\} \leq s$$

from n linear noisy measurements

$$y = Ax + w \quad \text{with } A \in \mathbb{R}^{n \times N} \quad \text{where } n \ll N.$$

Recovery via noise-constrained ℓ_1-minimization

$$\min_z \|z\|_1 \quad \text{subject to } \|Az - y\|_2 \leq \eta.$$

Implication for compressive sensing

Compressive sensing problem:
Recover an (approximately) s-sparse vector $x \in \mathbb{R}^N$, i.e.,

$$\|x\|_0 = \#\{\ell : x_\ell \neq 0\} \leq s$$

from n linear noisy measurements

$$y = Ax + w \quad \text{with } A \in \mathbb{R}^{n \times N} \quad \text{where } n \ll N.$$

Recovery via noise-constrained ℓ_1-minimization

$$\min_z \|z\|_1 \quad \text{subject to } \|Az - y\|_2 \leq \eta.$$

If $\|w\|_2 \leq \eta$ for known η and A is a draw of a Gaussian random matrix with $n \sim s \log(eN/s)$ then with high probability the minimizer $x^\sharp$ satisfies

$$\|x - x^\sharp\|_1 \lesssim \sigma_s(x)_1 + \eta\sqrt{\frac{s}{n}},$$

where $\sigma_s(x)_1 = \inf_{z:\|z\|_0 \leq s} \|x - z\|_1$.

Noise-blind compressive sensing

What if a good upper bound η for the noise vector w is not known?

Noise-blind compressive sensing

What if a good upper bound η for the noise vector w is not known?

Recovery of x from $y = Ax + w$ via equality-constrained ℓ_1-minimization

$$\min_z \|z\|_1 \quad \text{subject to } Az = y. \tag{2}$$

Noise-blind compressive sensing

What if a good upper bound η for the noise vector w is not known?

Recovery of x from $y = Ax + w$ via equality-constrained ℓ_1-minimization

$$\min_z \|z\|_1 \quad \text{subject to } Az = y. \tag{2}$$

Analysis based on

- Null space property of A: For some $\rho \in (0,1)$ and for all $v \in \ker A \setminus \{0\}$ and all $S \subset \{1, \ldots, N\}$, $\#S = s$

  $$\|v_S\|_1 \leq \rho \|v_{S^c}\|_1.$$

  Satisfied for many random matrices if $n \sim \rho^{-2} s \log(eN/s)$.
- ℓ_1-quotient property of A with respect to norm $|||w|||$: For every $w \in \mathbb{R}^n$ there exists a vector $v \in \mathbb{R}^N$ such that $Av = w$ and

  $$\|v\|_1 \leq L^{-1} \, |||w|||.$$

If A satisfies both properties then minimizer $x^\sharp$ of (2) satisfies

$$\|x^\sharp - x\|_1 \lesssim \sigma_s(x)_1 + |||w|||$$

The ℓ_1-quotient property for random matrices

Let $|||\cdot|||$ be the gauge norm associated to $(K_p(X))^\circ$, i.e.,

$$|||x|||_p = \inf\{t > 0 : x \in t(K_p(X))^\circ\}.$$

Set $A = (X_1| \cdots |X_N)$ where $X_1, \ldots, X_N$ are independent copies of a random vector X satisfying our (weak) assumptions.

Main result: A satisfies ℓ_1-quotient property w.r.t. $|||\cdot|||_p$ for $p = \alpha \log(eN/n)$ with probability at least $1 - 2\exp(-cN^{1-\alpha} n^\alpha)$.

The ℓ_1-quotient property for random matrices

Let $\|\!|\cdot|\!\|$ be the gauge norm associated to $(K_p(X))^\circ$, i.e.,

$$\|\!|x|\!\|_p = \inf\{t > 0 : x \in t(K_p(X))^\circ\}.$$

Set $A = (X_1|\cdots|X_N)$ where $X_1, \ldots, X_N$ are independent copies of a random vector X satisfying our (weak) assumptions.

Main result: A satisfies ℓ_1-quotient property w.r.t. $\|\!|\cdot|\!\|_p$ for $p = \alpha \log(eN/n)$ with probability at least $1 - 2\exp(-cN^{1-\alpha} n^\alpha)$.

Example: If X is a standard Gaussian random vector or a vector with independent student t-distributed random variables with $d = 2\log(N)$ degrees of freedom then for $y = Ax + w$ the minimizer $x^\sharp$ of equality-constrained ℓ_1-minimization satisfies

$$\|x^\sharp - x\|_1 \lesssim \sigma_s(x)_1 + \sqrt{\log(eN/n)}\|w\|_2 \sim \sigma_s(x)_1 + \|w\|_2\sqrt{\frac{s}{n}}.$$

Error bound involves true noise-level $\|w\|_2$ instead of upper estimate η and reconstruction procedure does not require any knowledge about $\|w\|_2$.

Proof method of main theorem

Mendelson's small ball method:

- By duality need to prove a high probability lower bound

$$\inf_{t \in \partial K_p(X)} \|\Gamma t\|_\infty \geq c, \quad \text{where } \Gamma = (X_1 | \ldots, X_N)^T.$$

Proof method of main theorem

Mendelson's small ball method:

- By duality need to prove a high probability lower bound

$$\inf_{t \in \partial K_p(X)} \|\Gamma t\|_\infty \geq c, \qquad \text{where } \Gamma = (X_1 | \ldots, X_N)^T.$$

- Show that with high probability the number $\#\{i : \langle X_i, t \rangle \geq 1\}$ is high for each individual t.

Proof method of main theorem

Mendelson's small ball method:

- By duality need to prove a high probability lower bound

$$\inf_{t \in \partial K_p(X)} \|\Gamma t\|_\infty \geq c, \quad \text{where } \Gamma = (X_1 | \ldots, X_N)^T.$$

- Show that with high probability the number $\#\{i : \langle X_i, t \rangle \geq 1\}$ is high for each individual t.
- Extend to a net of $\partial K_p(X)$ by union bound
- Uniformly bound local variations $\#\{i : |\langle X_i, t - \pi t \rangle| \geq 1/2\}$

Thanks very much for your attention!