Super-Gaussian directions of random vectors

Bo'az Klartag

Weizmann Institute & Tel Aviv University

IMU-INdAM Conference in Analysis,

Tel Aviv, June 2017

Gaussian approximation

- Many distributions in $\mathbb{R}^n$, for large n, have approximately Gaussian marginals.

Classical central limit theorem

Let $X = (X_1, \ldots, X_n)$ be a random vector with i.i.d coordinates, finite third moment. Then for $\theta = (1, \ldots, 1)/\sqrt{n}$, the random variable

$$\langle \theta, X \rangle = \sum_i \theta_i X_i$$

is approx. Gaussian (Kolmogorov distance $\leq C\mathbb{E}|X_1|^3/\sqrt{n}$).

- If the coordinates are not identically-distributed, but still independent, can take another $\theta \in S^{n-1}$.
- Geometric interpretation: approx. gaussian directions.
- We may replace independence by some weak dependence.

Marginals of high-dimensional distributions

- Maxwell's principle: If X is uniformly distributed in a Euclidean ball, then $\langle X, \theta \rangle$ is approx. Gaussian.

Theorem (CLT for convex sets, K. '07, Fleury '09, Guédon-Milman '11, Lee-Vempala '16)

If X is uniformly distributed in some convex domain in $\mathbb{R}^n$, then for some $\theta \in S^{n-1}$, the random variable

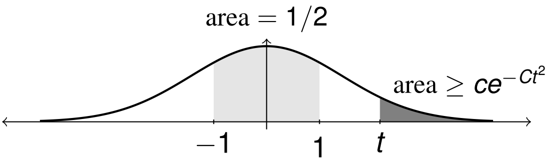
$$\langle X, \theta \rangle$$

is approx. Gaussian.
(Kolmogorov distance $\leq C/n^{\alpha}$. Best α unknown, at least $\approx 1/4$)

- However, if $X = (X_1, \ldots, X_n)$ has independent coordinates, Cauchy-distributed (density $t \mapsto C_\gamma/(\gamma^2 + t^2)$), then all marginals are Cauchy as well, far from Gaussian!
- Similarly, α-stable for $0 < \alpha < 2$.

Universality in high dimensions?

- Different random vectors in high dimension may have very different marginals. Still, the Cauchy distribution and all of the other α-stable distributions are *super-Gaussian*:



Definition ("A centered super-Gaussian random variable")

A random variable Y is **super-Gaussian of length $L > 0$** with parameters $\alpha, \beta > 0$ if $\mathbb{P}(Y = 0) = 0$ and for any $0 \leq t \leq L$,

$$\mathbb{P}\left(\frac{Y}{\sigma} \geq t\right) \geq \alpha e^{-t^2/\beta} \quad \text{and} \quad \mathbb{P}\left(\frac{Y}{\sigma} \leq -t\right) \geq \alpha e^{-t^2/\beta},$$

where $\sigma = \mathrm{Median}(|Y|)$ is any median of $|Y|$.

A modest conjecture

- In all of the examples of random vectors $X \in \mathbb{R}^n$ above, for some $\theta \in S^{n-1}$, the random variable
$$\langle X, \theta \rangle$$
has a visible tail, for many standard deviations.

Example

When X is distributed uniformly in a centered Euclidean ball in $\mathbb{R}^n$, for any $\theta \in S^{n-1}$, the density of $\langle X, \theta \rangle$ is proportional to

$$t \mapsto \left(1 - \frac{t^2}{A^2 n}\right)_+^{(n-1)/2} \approx e^{-t^2/(2A^2)}.$$

Thus $\langle X, \theta \rangle$ is super-Gaussian of length $c\sqrt{n}$, and not longer (with parameters c_1, c_2 that are universal constants).

Main result

There are always super-Gaussian directions, of length $c\sqrt{n}$:

Theorem (K., '15)

Let X be a random vector **with density** *in $\mathbb{R}^n$. Then there exists a fixed vector $\theta \in S^{n-1}$ such that $\langle X, \theta \rangle$ is super-Gaussian of length $c_1\sqrt{n}$ with parameters $c_2, c_3 > 0$.*

- Here, $c_1, c_2, c_3 > 0$ are universal constants, independent of the density of X and of the dimension.
- Optimal up to constants, as shown by the Euclidean ball.

Why do we need a density?
(cannot take a deterministic random vector, for instance).

- When X is distributed uniformly in a convex set, proven by Pivovarov '10 (in the unconditional case, up to log) and by Paouris '12 (under hyperplane conjecture). Previous logarithmic estimate in K. '10.

Discrete random vectors

Definition

Let X be a random vector in $\mathbb{R}^n$, let $0 < d \leq n$. The **effective rank** of X is at least d if for any linear subspace $E \subseteq \mathbb{R}^n$,

$$\mathbb{P}(X \in E) \leq \dim(E)/d,$$

with equality iff $\exists F \subseteq \mathbb{R}^n$ with $E \oplus F = \mathbb{R}^n$, $\mathbb{P}(X \in E \cup F) = 1$.

Examples of random vectors whose effective rank is exactly n

1. A random vector with density in $\mathbb{R}^n$.
2. A random vector that is distributed uniformly on a finite set that spans $\mathbb{R}^n$ and does not contain the origin.
3. The cone volume measure of any convex body in $\mathbb{R}^n$ with barycenter at the origin (Böröczky, Lutwak, Yang, and Zhang '15, Henk and Linke '14)

Even less assumptions

Theorem (K. '15)

Let $d \geq 1$ and let X be a random vector in a finite-dimensional linear space, whose effective rank is at least d.

Then there exists a non-zero, fixed, linear functional ℓ such that the random variable $\ell(X)$ is super-Gaussian of length $c_1\sqrt{d}$ with parameters $c_2, c_3 > 0$.

- We cannot assert that most directions are super-Gaussian.

The simplest example

Suppose $\mathbb{P}(X = e_i) = 1/n$ for $i = 1, \ldots, n$. Then for a **typical** $\theta \in S^{n-1}$,

$$\langle X, \theta \rangle$$

is approx. Gaussian, and is super-Gaussian of length $c\sqrt{\log n}$. However, we get length $c\sqrt{n}$ in the direction of $(\theta + e_1)/|\theta + e_1|$.

Angularly isotropic position

Definition

A random vector X in $\mathbb{R}^n$ with $\mathbb{P}(X = 0) = 0$ is *angularly isotropic* if

$$\mathbb{E} \left\langle \frac{X}{|X|}, \theta \right\rangle^2 = \frac{1}{n} \qquad \forall \theta \in S^{n-1}.$$

- The condition that $\mathbb{P}(X \in E) \leq \dim(E)/n$ for any subspace E is necessary: Setting $\tilde{X} = X/|X|$,

$$\mathbb{P}(X \in E) = \mathbb{P}(\tilde{X} \in E) \leq \mathbb{E}|Proj_E \tilde{X}|^2 = \dim(E)/n.$$

Theorem (K. '10, BLYZ '15)

Any random vector with effective dimension at least n has a linear image which is angularly isotropic.

Proof of the main result – The range $[0, \sqrt{\log n}]$

Proposition

Assume X is angularly isotropic. Then for a random $\theta \in S^{n-1}$, with high probability

$$Y = \langle X, \theta \rangle$$

is super-Gaussian of shorter length, about $c\sqrt{\log n}$.

Proof idea: Let $X_1, \ldots, X_k$ be i.i.d copies of X with $k = \lfloor n^{1/10} \rfloor$. With high prob., these are k approximately-orthogonal vectors.

- Therefore the "simplest example" analysis applies.

For a typical direction $\theta \in S^{n-1}$, the numbers

$$\langle X_1, \theta \rangle, \ldots, \langle X_k, \theta \rangle$$

look like a Gaussian sample. We reach roughly $\sqrt{\log n}$ standard deviations.

Sudakov minoration

- In order to deal with the range $t \gg \sqrt{\log n}$, we shall use Sudakov's theorem.

Theorem (Sudakov, 1969)

Let $N \geq 1, \alpha > 0$ and let $x_1, \ldots, x_N \in \mathbb{R}^n$. Assume that

$$|x_i - x_j| \geq \alpha \qquad \text{for any } i \neq j.$$

Let $\Theta \in S^{n-1}$ be a random vector, distributed uniformly. Then,

$$\mathbb{E} \max_{i=1,\ldots,N} \langle x_i, \Theta \rangle \geq c\alpha \sqrt{\frac{\log N}{n}},$$

- We would get roughly the same estimate if the random variables $\langle x_i, \Theta \rangle$ were independent with $|x_i| = \alpha$ for all i.

- Why is the "simplest example" stuck at length $\sqrt{\log n}$? Because some small cap $B \subset S^{n-1}$ has too large a mass.

Set $M = \mathrm{Median}(|X|)$ and let $t_0 \geq 0$ be defined via

$$\exp(-t_0^2) = \sup_{v \in S^{n-1}} \mathbb{P}\left(|X| \geq M \text{ and } \left| \frac{X}{|X|} - v \right| \leq \frac{1}{2} \right).$$

- Note that $t_0 \lesssim \sqrt{n}$. By angular isotropicity, $t_0 \gtrsim \sqrt{\log n}$.

Proposition

Assume X is angularly isotropic. Then for a random $\theta \in S^{n-1}$, w.h.p $\langle X, \theta \rangle$ is super-Gaussian of length at least ct_0.

Proof: Let $X_1, \ldots, X_k$ be i.i.d copies of X with $k = \lfloor e^{t_0^2/4} \rfloor$. With high prob., $\{X_i/|X_i|\}$ are $1/2$-separated (after trimming). Now use Sudakov's minoration and measure concentration. $\square$

Proof ideas – the range $[t_0, \sqrt{n}]$

Let $B(v, 1/2)$ be a fixed cap of mass $\exp(-t_0^2)$.

A simple modification

Replace the random direction $\theta \in S^{n-1}$ by $\eta := (\theta + v)/|\theta + v|$.

- Well, the random vector $\eta \in S^{n-1}$ is not distributed uniformly on S^{n-1}. Still, previous analysis applies.

What have we obtained so far?

For a typical choice of η, we have $\mathrm{Median}(|\langle X, \eta\rangle|) \sim M/\sqrt{n}$ and

$$\mathbb{P}\left(\langle X, \eta\rangle \geq \frac{tM}{\sqrt{n}}\right) \geq c\exp(-Ct^2) \qquad \forall 0 \leq t \leq t_0.$$

- Since η is biased towards v, then for $t_0 < t < \sqrt{n}/5$,

$$\mathbb{P}\left(\langle X, \eta\rangle \geq \frac{tM}{\sqrt{n}}\right) \geq \mathbb{P}\left(\langle X, \eta\rangle \geq \frac{M}{5}\right) \geq ce^{-Ct_0^2} \geq ce^{-Ct^2}. \quad \square$$

The end

Thank you!