A Central Limit Theorem for Convex Bodies

Bo'az Klartag

Princeton University

A lecture at the educational workshop:

"Convex geometry - Analytic aspects"

Cortona, Italy
June, 2007.

We are given an arbitrary convex body $K \subset \mathbb{R}^n$. The dimension n is assumed very large.

- Sample a random point from the body K. That is, the random vector X is disributed uniformly in K.

Is there anything interesting to say about X? Have a look at the coordinates $X = (X_1, \dots, X_n)$. What is their distributions?

Well, this obviously depends on K. Examples:

(i) K is a Euclidean ball. Then $X_1, \dots, X_n$ are identically-distributed, with the density

$$t \mapsto c_n \left(1 - \frac{t^2}{n}\right)^{\frac{n-1}{2}} \approx \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

(we normalized the Euclidean ball)

For the Euclidean ball, all marginals are approximately gaussian. (was known to Maxwell)

- The gaussian approximation holds also if we linearly change the coordinate system.

Even for an almost degenrate ellipsoid, all of the marginals are approximately normal (with various means / variances).

A second example, easy to compute:

(ii) $K = [0,1]^n$ is a cube. Then $X_1, \ldots, X_n$ are independent, and are all distributed uniformly in $[0,1]$.

This seems quite different from the Euclidean ball.

However, we used very special coordinates. What happens when we view X with respect to a more "generic" system of coordinates?

Suppose we look at the from another angle. (equivalently, we rotate the cube).

Let $Y_1, \ldots, Y_n$ be the coordinates of the random point X with respect to a Walsh-Hadamard matrix (say, the dimension is a power of two). Then,

$$Y_i = \frac{\pm X_1 + \ldots + \pm X_n}{\sqrt{n}}$$

with different choice of signs for each i.

From the **classical central limit theorem**:

- All $Y_1, \ldots, Y_n$ are distributed approximately according to the gaussian law.

Actually, with respect to "most" orthonormal bases, the coordinates are approximately gaussian.

Recall: $K \subset \mathbb{R}^n$ a convex body, n very large. The random vector X is uniform in K.

In the examples of the ball and the cube, we saw that with respect to a "typical" choice of an orthonormal basis, all of the coordinates were approximately normal.

- This is a general phenomenon.

<u>Main Theorem:</u> Suppose K is an arbitrary n-dimensional convex body. With respect to an appropriate (or "generic") basis,

- The coordinates $X_1, \ldots, X_n$ of the random vector $X \in K$ are approximately gaussian.

- Moreover, $X_{n_1}, \ldots, X_{n_\ell}$ are approximately independent, for any distinct $n_1, \ldots, n_\ell$, when $\ell \leq c n^\kappa$.

Here, $c, \kappa > 0$ are universal constants.

What exactly do we mean by "approximately"?

The total variation distance between two random variables X and Y is

$$d_{TV}(X,Y) = \sup_A |\mathbb{P}\{X \in A\} - \mathbb{P}\{Y \in A\}|$$

where sup runs over all measurable A.

- "X_i is approximately gaussian" means that

$$d_{TV}(X_i, \Gamma) \leq C/n^\kappa$$

With Γ being normal, and, say, $\kappa \approx 1/20$.

- "$X_{n_1}, \ldots, X_{n_\ell}$ are approx. indep." means

$$d_{TV}\left[(X_{n_1}, \ldots, X_{n_\ell}), \Gamma_\ell\right] \leq C/n^\kappa$$

where Γ_ℓ is an ℓ-dim. gaussian.

What is "generic"? When $\mathbb{E}X = 0, Cov(X) = I$, then a random orthonormal basis $\{v_1, \ldots, v_n\} \in O(n)$ will do.

Recent History

The “modern history” of this theorem begins with two (independent) papers:

- Anttila, Ball and Perissinaki ’03
- Brehm and Voigt ’00

It was conjectured in these articles that, typical marginals of arbitrary convex bodies are approx. gaussian.

Important cases were established (e.g., ABP considered certain uniformly convex bodies).

Following BV and ABP, significant contributions by: Bastero, Bernués, Bobkov, Hinow, Koldobsky, Lifshits, E. and M. Meckes, E. Milman, Naor, Paouris, Romik, S. Sodin, Vogt, Wojtaszczyk and others.

Less recent History

A beautiful theorem of Sudakov '78, with no convexity involved.

We follow the presentation of Bobkov '03. Assume that X is any random vector in $\mathbb{R}^n$ with

$$\mathbb{E}X = 0, \quad Cov(X) = Id$$

(i.e., $\mathbb{E}X_iX_j = \delta_{i,j}$).

<u>**Theorem:**</u> (Sudakov '78)
There exists a (real valued) random variable Z, depending on X, such that for most $\theta \in S^{n-1}$,

$$d_{Lévy}(\langle X, \theta \rangle \, , \, Z \,) \leq \varepsilon_n$$

where $\varepsilon_n \searrow 0$ as $n \to \infty$ (independently of X). Say, $\varepsilon_n \leq Cn^{-1/5}$.

Here, $d_{Lévy}(X, Y)$ is the minimal $\delta > 0$ s.t.

$$\mathbb{P}\{X \leq t - \delta\} - \delta \leq \mathbb{P}\{Y \leq t\} \leq \mathbb{P}\{X \leq t + \delta\} + \delta$$

for all $t \in \mathbb{R}$.

What is that mysterious common distribution?

A random vector X in $\mathbb{R}^n$ is isotropic if

$$\mathbb{E}X = 0, \quad Cov(X) = I.$$

This just means normalization.

Theorem (Sudakov '78): For any isotropic random vector X in $\mathbb{R}^n$, most of its marginals are distributed approximately

$$|X| \cdot \frac{\Gamma}{\sqrt{n}}$$

where $\Gamma \sim N(0,1)$ is a standard normal random variable, independent of X.

Corollary (Diaconis-Freedman '84):
Suppose X is r.v. in $\mathbb{R}^n$, for large n, with $\mathbb{E}X = 0, Cov(X) = Id$. Assume that

$$\mathbb{P}\left(\left| \frac{|X|}{\sqrt{n}} - 1 \right| > \varepsilon \right) < \varepsilon$$

with $\varepsilon << 1$. Then most of the marginals of X are approximately gaussian.

Also related: Gromov '88, von Weizsäcker '97.

How come such a general, strong principle, such as Sudakov's theorem be true, with almost no assumptions at all?

• The driving force is the "concentration of measure" phenomenon in high dimensions.

The proof is conceptually very simple and illuminating. It goes as follows:

Suppose X is an isotropic random vector in $\mathbb{R}^n$. Then,

Mostly almost-gaussian marginals

$$\Updownarrow$$

Most of the mass in a thin spherical shell

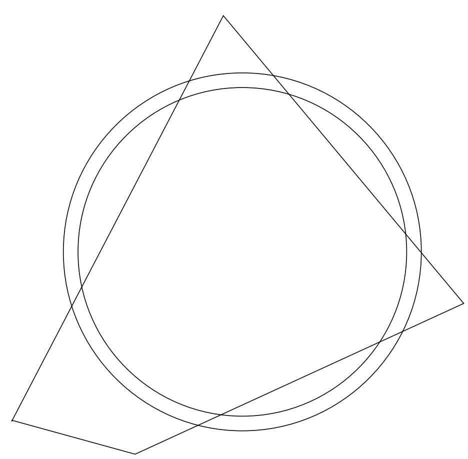
Back to convex sets: We need to prove that most of their mass lies in a thin spherical shell.

Example: Consider the density

$$\frac{1}{2} \cdot \frac{1}{(2\pi)^{n/2}} e^{-|x|^2/2} \; + \; \frac{1}{2} \cdot \frac{1}{(4\pi)^{n/2}} e^{-|x|^2/4}$$

on $\mathbb{R}^n$. None of its marginals are close to gaussian.

Most of the mass of an isotropic convex body is in a thin spherical shell.



Why Convex Sets?

we have the classical Brunn-Minkowski inequality '87:

$$Vol(A+B)^{1/n} \geq Vol(A)^{1/n} + Vol(B)^{1/n}$$

for any measurable sets $A, B \subset \mathbb{R}^n$.

The Brunn-Minkowski inequality implies that marginals of convex sets have log-concave densities.

Recall that a function $f : \mathbb{R}^n \to [0, \infty)$ is log-concave if

$$f(\lambda x + (1-\lambda)y) \geq f(x)^\lambda f(y)^{1-\lambda}$$

for all $x, y \in \mathbb{R}^n$ and $0 < \lambda < 1$. The characteristic function of a convex set is log-concave.

- If X has a log-concave density, then also $Proj_E(X)$ has a log-concave density, for any subspace E. [Prékopa-Leindler '73]

Sketch of the Proof

We are given a random vector X in $\mathbb{R}^n$ with

$$\mathbb{E}X = 0, Cov(X) = I$$

, whose density is log-concave.

Concentration Phenomenon:
Suppose $E \subset \mathbb{R}^n$ is a random, k-dimensional subspace (for certain k). Then, with large probability,

The density of $Proj_E(X)$ is approx.
spherically-symmetric (in L^∞-sense).

Gromov '88 remarked that $Proj_E(X)$ is approx. radial, in a weak sense, even without the log-concavity assumption. Similar to Sudakov.

By Brunn-Minkowski (or Prékopa-Leindler),

The density of $Proj_E(X)$ is log-concave.

To summarize, for a random subspace $E \subset \mathbb{R}^n$, the density f_E of $Proj_E(X)$ is both

(i) Almost radial,

(ii) Log-concave.

<u>**Lemma**</u> (Laplace asymptotic method): Let $p > 0$, let $f : [0,\infty) \to [0,\infty)$ be log-concave and integrable. Then there exists $r > 0$ with

$$\int_{r(1-\varepsilon)}^{r(1+\varepsilon)} t^p f(t)dt \geq \left(1 - Ce^{-c\varepsilon^2 p}\right) \int_0^\infty t^p f(t)dt$$

for all $0 < \varepsilon < 1$.

Suppose the density f_E is exactly spherically-symmetric. Then,

$$\mathbb{P}\{ \ (1-\varepsilon)r \leq |Proj_E(X)| \leq (1+\varepsilon)r \ \}$$

$$= \int_{S(E)} \int_{r(1-\varepsilon)}^{r(1+\varepsilon)} t^{k-1} f_E(t\theta)dtd\theta \geq 1 - Ce^{-c\varepsilon^2 k}.$$

Uniform distribution on the sphere

Consequently, for a typical subspace $E \subset \mathbb{R}^n$ (of dimension k),

- The random vector $Proj_E(X)$ is approx. uniform on a sphere of radius r.

We apply now a further projection, to any subspace $F \subset E$ of dimension $\varepsilon^2 k$. By Maxwell's principle,

$$Proj_F(X) = Proj_F(Proj_E(X))$$

is ε-close to gaussian, in the total variation distance.

Thus, a random projection of X – an arbitrary log-concave, isotropic random vector – is approx. gaussian.

This completes the proof, for all log-concave densities. Regarding the "thin shell" with some extra effort we get

<u>**Theorem:**</u> For any $0 < \varepsilon < 1$,

$$\mathbb{P}\left\{\left|\frac{|X|}{\sqrt{n}} - 1\right| \geq \varepsilon\right\} \leq C\,\mathrm{exp}(\,-c\,\varepsilon^{\alpha} n^{\beta}\,)$$

where, say, $\alpha = 3.33$ and $\beta = 0.33$.

So the "thin shell" that contains 99% of the mass, has width at most $\approx Cn^{-1/10}$, relative to the radius. Probably non-optimal.

In the large deviations scale, a sharp estimate is known.

<u>**Theorem:**</u> (Paouris '06) For any $t > 1$,

$$\mathbb{P}\{|X| \geq Ct\,\mathbb{E}|X|\} \leq \mathrm{exp}(\,-t\sqrt{n}\,)$$

where $C > 0$ is a universal constant.

The classical CLT tells us that $X_1 + \ldots + X_n$ is approximately gaussian. For arbitrary convex sets, it's certainly false.

Suppose that our log-concave density $f : \mathbb{R}^n \to [0, \infty)$ is "unconditional":

$$f(x_1, ..., x_n) = f(|x_1|, ..., |x_n|) \qquad \forall x \in \mathbb{R}^n.$$

__**Theorem:**__ Suppose X is an isotropic random vector, with an unconditional, log-concave density. Then, for any $t \in \mathbb{R}$,

$$\left| \mathbb{P}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i \le t \right) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-s^2/2} ds \right| \le \frac{C}{\sqrt{n}}$$

and more generally, for any $(\theta_1, \ldots, \theta_n) \in S^{n-1}$,

$$\left| \mathbb{P}\left(\sum_{i=1}^n \theta_i X_i \le t \right) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-s^2/2} ds \right| \le C \sqrt{\sum_{i=1}^n \theta_i^4}.$$

In the unconditional case, we have a precise $C/\sqrt{n}$ estimate for thin shell.

Uniform Large Deviation Estimates

Suppose X is uniform in a convex body in $\mathbb{R}^n$. A classical fact (follows from Brunn-Minkowski):

<u>**Theorem**</u> (Borell '74): For any linear functional $\varphi : \mathbb{R}^n \to \mathbb{R}$ and $t \geq 1$,

$$\mathbb{P}\{|\varphi(X)| \geq t\,\mathbb{E}|\varphi(X)|\} \leq \exp(-ct)$$

where $c > 0$ is a universal constant.

A uniformly subexponential tail. This is sharp, as shown by the example of a truncated cone.

Suppose X is uniform in a centered ellipsoid. Then, for all linear functionals φ,

$$\mathbb{P}\{|\varphi(X)| \geq t\,\mathbb{E}|\varphi(X)|\} \leq \exp\left(-ct^2\right).$$

Moreover, the tail is very close to being gaussian.

Question: Is it true that for any convex body, there is a linear functional with a uniformly sub-gaussian tail?

(if true, a convex body cannot display "cone-type" behavior in all directions)

True for unconditional convex bodies (Bobkov-Nazarov '03) and for zonoids (Paouris '03). For arbitrary convex bodies:

Theorem (K. '05, simplified and slightly improved by Giannopoulos, Pajor, Paouris '06):

Suppose X is uniform in a convex set. Then there exists a non-zero linear functional $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that for any $t \geq 1$,

$$\mathbb{P}\{|\varphi(X)| \geq t\,\mathbb{E}|\varphi(X)|\} \leq \exp\left(-c\frac{t^2}{\log^2(t+1)}\right)$$

where $c > 0$ is a universal constant.