Rate of convergence in the central limit theorem for convex bodies

Bo'az Klartag
Princeton University

Samos, Greece
June 2007

Central Limit Theorem for Convex Bodies

- Let $X = (X_1, \ldots, X_n)$ be a random vector, uniformly distributed in a convex body $K \subset \mathbb{R}^n$.

  The dimension n is assumed very large.

- In some situations (e.g., K is the cube or the Euclidean ball) the random variable

$$\frac{X_1 + \ldots + X_n}{\sqrt{n}}$$

  is approximately gaussian.

Central Limit Theorem for Convex Bodies

- Let $X = (X_1, \ldots, X_n)$ be a random vector, uniformly distributed in a convex body $K \subset \mathbb{R}^n$.

  The dimension n is assumed very large.

- In some situations (e.g., K is the cube or the Euclidean ball) the random variable

$$\frac{X_1 + \ldots + X_n}{\sqrt{n}}$$

  is approximately gaussian.

- The gaussian approximation actually holds for all convex bodies, once we select an appropriate coordinate system.

Central Limit Theorem for Convex Bodies

- Normalize the random vector $X = (X_1, \ldots, X_n)$: Assume that

  $$\mathbb{E}X_i = 0, \qquad \mathbb{E}X_iX_j = \delta_{i,j}.$$

  We say that X is *normalized* or *isotropic*.

- **General Theorem**: Suppose X is an isotropic random vector, uniformly distributed in some convex body in $\mathbb{R}^n$. Then with respect to an appropriate (or "generic") orthogonal basis,
  - The random variables $X_1, \ldots, X_n$ are approximately gaussian.
  - The random variables $X_1, \ldots, X_n$ are approximately independent in ℓ-tuples ($\ell \leq n^c$).

Quantitative Information

Remarks on the above theorem:

- "approximately gaussian" means, say, *total variation distance*:

$$\sup_{A \subset \mathbb{R}} \left| \mathbb{P}\{X_i \in A\} - \frac{1}{\sqrt{2\pi}} \int_A e^{-t^2/2} dt \right| \leq \frac{C}{n^c}.$$

- "approximately independent in ℓ-tuples" means that the joint distribution is close to the product of the 1D distributions.
- "generic orthogonal basis" means that a random basis $\{v_1, \ldots, v_n\} \in O(n)$ works, with probability $> 1 - \exp(-\sqrt{n})$.

This *central limit theorem for convex bodies* was explicitly conjectured by Anttila-Ball-Perissinaki '03 and Brehm-Voigt '00.

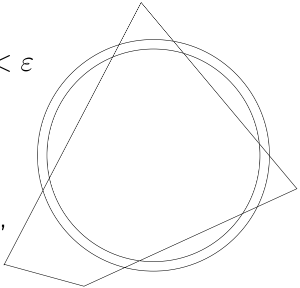
Additional Contributions: Bastero, Bernués, Bobkov, Hinow, Koldobsky, Lifshits, E. and M. Meckes, E. Milman, Naor, Paouris, Romik, S. Sodin, Vogt, Wojtaszczyk and others

Thin Spherical Shell

This *central limit theorem for convex bodies* follows from a "spherical thin shell" estimate:

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

for $\varepsilon << 1$.



- The equivalence of a "thin shell estimate" and "central limit theorems" is classical, due to Sudakov '78, Diaconis-Freedman '84, Anttila-Ball-Perissinaki '03 and others.
- Its proof has two ingredients: The concentration of measure phenomenon, and Maxwell's principle, stating that marginals of the Euclidean sphere are approximately gaussian.

Thin Spherical Shell

Thus, to prove our main theorem, it suffices to show that

$$\mathbb{E}\left(\frac{|X|}{\sqrt{n}} - 1\right)^2 << 1.$$

- First proof: An upper bound of $\frac{C}{\log n} << 1$. This results in weaker, logarithmic estimates in place of Cn^{-c}.
- Another proof, by Fleury, Guédon and Paouris, yielded similar logarithmic bounds.
- Yet another proof gives a bound

$$\mathbb{E}\left(\frac{|X|}{\sqrt{n}} - 1\right)^2 \leq \frac{C}{n^c}$$

with, say, $c \approx 1/5$.

Sketch of Proof

A density function in $\mathbb{R}^n$ is *log-concave* if it takes the form e^{-H} with $H : \mathbb{R}^n \to (-\infty, \infty]$, a convex function.

- When X is uniformly distributed in a convex set, its density is obviously log-concave.
- When X, Y are independent with log-concave densities, then also $X + Y$ and $Proj_E(X)$ have log-concave densities (Prékopa-Leindler inequality).

Observation: Suppose X is an isotropic random vector, whose density f is log-concave and radial. Then,

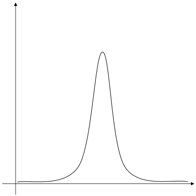
$$\mathbb{E}\left(\frac{|X|}{\sqrt{n}} - 1\right)^2 \leq \frac{C}{n}.$$

Radial, log-concave densities

Explanation of the observation: The density of the real-valued random variable $|X|$ is

$$t \mapsto t^{n-1} f(t)$$

with f log-concave.



Such densities are necessarily very peaked:

$$\text{``} \quad t^{n-1} f(t) \;\leq\; t_0^{n-1} f(t_0) \exp\left(-Cn(t-t_0)^2\right) \quad \text{''}$$

where $t_0 \approx \sqrt{n}$ is the point where $t \mapsto t^{n-1} f(t)$ attains its maximum.

It follows that

$$\mathbb{E}\left(\frac{|X|}{\sqrt{n}} - 1\right)^2 \leq \frac{C}{n}.$$

Sketch of Proof

Problem: The density of our random vector X is assumed to be log-concave, but not at all radial.

We would like to "transform" our random vector somehow, and then to apply the result for radial & log-concave distributions.

Two simple observations:

- It is enough to prove that

$$\mathbb{E}\left(\frac{|X+Y|}{\sqrt{2n}} - 1\right)^2 \leq \frac{1}{n^c}$$

where Y is an independent, isotropic vector (say, gaussian).

- Suppose E is a random k-dimensional subspace in $\mathbb{R}^n$, independent of everything. Then, with high probability,

$$|Proj_E(X+Y)| \approx \sqrt{\frac{k}{n}}|X+Y|.$$

Sketch of Proof

Therefore, in order to show that $|X|$ is concentrated, it suffices to prove that the real-valued random variable

$$|Proj_E(X+Y)|$$

is concentrated, where Y is an independent gaussian, and E an independent random subspace.

Claim: For most choices of E, the random vector $Proj_E(X+Y)$ is log-concave and approximately radial.

- If the claim is true, and the approximation is good enough (e.g., total-variation), then we may conclude by applying the result on radial & log-concave densities.

Lipschitz Functions in High Dimension

The density of $Proj_E(X+Y)$ is always log-concave, by Prékopa-Leindler.

- How can we show that $Proj_E(X+Y)$ is approximately radial?

Denote by $f_E : E \to [0,\infty)$ the density of $Proj_E(X+Y)$.
We hope that $\log f_E(x)$ is Lipschitz in $E \in G_{n,k}$ and $x \in E$.

In other words, fix $E_0 \in G_{n,k}$ and $x_0 \in E_0$, and consider the map

$$M(U) = \log f_{U(E_0)}(U(x_0)) \qquad (U \in SO(n)).$$

We need to show that M is Lipschitz. Then we could use concentration inequalities.

Concentration of Measure

Some analysis with log-concave densities shows that under mild assumptions, for $U_1, U_2 \in SO(n)$,

$$|M(U_1) - M(U_2)| \leq Ck^2 |U_1 - U_2|_{HS},$$

the Hilbert-Schmidt norm. (The power of k should be improved.)

We make an essential use of the convolution with the gaussian (i.e., $X + Y$ in place of X). Otherwise, M would not even be finite.

<u>**Theorem**</u>: [Gromov-Milman '83]
Select random, independent, $U_1, U_2 \in SO(n)$. Then,

$$\mathbb{P}\left\{|M(U_1) - M(U_2)| \geq \varepsilon\right\} \leq \exp\left(-cn\varepsilon^2/k^4\right).$$

Tying up loose ends

If $k << n^{1/4}$, then for a random $U \in SO(n)$,

$$\log f_{U(E_0)}(U(x_1)) \approx \log f_{U(E_0)}(U(x_2))$$

if $x_1, x_2 \in E_0$ with $|x_1| = |x_2|$.

Consequently, f_E is typically approximately radial pointwise, and hence also in total-variation.

This finishes the proof (up to details). We obtain

$$\mathbb{E}\left(\frac{|X|}{\sqrt{n}} - 1\right)^2 \approx Var\left(\frac{|X|}{\sqrt{n}}\right) \leq \frac{C}{n^c}.$$

Optimal exponents are still unknown in general.

Unconditional Convex Bodies

Consider an isotropic random vector X, with a log-concave density f, such that

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

i.e., unconditional densities. Then we can prove

$$\mathbb{E}\left(\frac{|X|}{\sqrt{n}} - 1\right)^2 \leq \frac{C}{n},$$

the optimal estimate. The methods are completely different.

- For simplicity, assume that X is uniform in a convex body $K \subset \mathbb{R}^n$. Let $\rho : \mathbb{R}^n \to \mathbb{R}$ be a convex function with

$$K = \{x \in \mathbb{R}^n; \rho(x) \leq 1\}.$$

Hörmander's Formula

Proposition: ["Hörmander's Formula"]
For any smooth function $u : K \to \mathbb{R}$ with $\partial_n u = 0$ on ∂K,

$$\int_K (\triangle u)^2 = \int_K |\nabla^2 u|_{HS}^2 + \int_{\partial K} \nabla^2 \rho(\nabla u, \nabla u).$$

- Since K is convex, then:

$$\int_K (\triangle u)^2 \geq \int_K |\nabla^2 u|_{HS}^2 = \sum_{i=1}^n \int_K |\nabla \partial^i u|^2.$$

- We dualize, and get that for any function $f : K \to \mathbb{R}$,

$$\int_K f = 0 \quad \Rightarrow \quad \int_K f^2 \leq \sum_{i=1}^n \|\partial^i f\|_{H^{-1}(K)}^2,$$

where $\|v\|_{H^{-1}(K)} = \sup\left\{\int_K uv;\ \int_K |\nabla u|^2 \leq 1\right\}$.

Transportation of Measure

For a given function f, how can we compute

$$\|f\|_{H^{-1}(K)} = \sup\left\{\int_K fu;\ \int_K |\nabla u|^2 \leq 1\right\}?$$

- The $H^{-1}(K)$-norm of f makes sense only when $\int_K f = 0$.

Denote by λ the restriction of the Lebesgue measure to K.

Proposition: [Brenier '87, Villani '02] When $\int_K f = 0$,

$$\|f\|_{H^{-1}(K)} \leq \liminf_{\varepsilon\to 0^+} \frac{W_2\left(\lambda, (1+\varepsilon f)\lambda\right)}{\varepsilon},$$

where W_2 is the L^2-Wasserstein distance between the measures.

(*Actually, up to some technicalities, it's an equality*)

Transportation of Measure

- Recall that for two measures μ and ν on $\mathbb{R}^n$,

$$W_2^2(\mu,\nu) = \inf_{\gamma} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x-y|^2 d\gamma(x,y)$$

  where the infimum runs over all couplings γ of μ and ν: That is, measures on $\mathbb{R}^n \times \mathbb{R}^n$ whose corresponding marginals are μ and ν.

- To summarize: For any function $f : K \to \mathbb{R}$ with $\int_K f = 0$,

$$\int_K f^2 \le \sum_{i=1}^n \|\partial^i f\|_{H^{-1}(K)}^2 \le \sum_{i=1}^n \liminf_{\varepsilon \to 0^+} \frac{W_2^2\left(\lambda, (1+\varepsilon \partial^i f)\lambda\right)}{\varepsilon^2}.$$

Unconditional Convex Bodies

- Consider the function $f(x) = |x|^2 - n$. Then,

$$\int_K \left(|x|^2 - n\right)^2 dx \leq \sum_{i=1}^{n} \liminf_{\varepsilon \to 0^+} \frac{W_2^2\left(\lambda, (1 + 2\varepsilon x_i)\lambda\right)}{\varepsilon^2}.$$

- When K is unconditional, direct analysis shows that

$$\frac{1}{Vol_n(K)} \liminf_{\varepsilon \to 0^+} \frac{W_2^2\left(\lambda, (1 + 2\varepsilon x_i)\lambda\right)}{\varepsilon^2} \leq \frac{8}{3}\mathbb{E}X_i^4 \leq 16,$$

where X is the isotropic random vector, uniform in K. Hence,

$$\mathbb{E}\left(|X|^2 - n\right)^2 \leq 16n \quad \Longrightarrow \quad \mathbb{E}\left(\frac{|X|}{\sqrt{n}} - 1\right)^2 \leq \frac{C}{n}.$$

The End

Thank you!