Phase Transition Phenomena in Integral Geometry

Martin Lotz

Warwick University

joint work with

Joel Tropp (Caltech)

and

Dennis Amelunxen, Michael McCoy, Ivan Nourdin, Giovanni Peccati

Jena, September 18, 2019

Motivation

Let $\boldsymbol{x}_0 \in \mathbb{R}^n$ be s-sparse, $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_0$ for $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ $(s < m < n)$.

$$\text{minimize} \quad \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}. \tag{$\star$}$$

When is $\boldsymbol{x}_0$ the unique solution of $(\star)$?

- Every row of $\boldsymbol{A}$ is seen as a measurement or observation that reveals information about $\boldsymbol{x}_0$.
- Motivation: applications in seismic imaging, signal processing, medical imaging, statistics and machine learning, ...

Compressed Sensing

Let $\boldsymbol{x}_0 \in \mathbb{R}^n$ be s-sparse, $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_0$ for random $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ $(s < m < n)$.

$$\text{minimize} \quad \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

- Donoho, Candès, Romberg & Tao, Rudelson & Vershynin: Successful recovery of $\boldsymbol{x}_0$ when

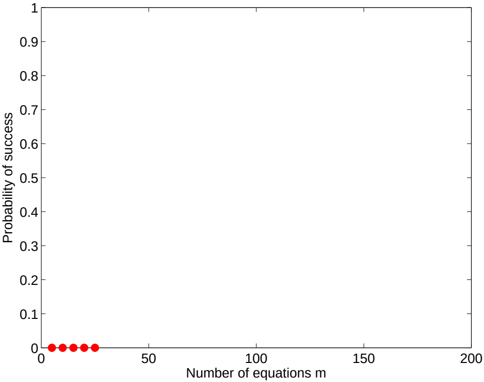
$$m \geq \mathrm{const} \cdot \log(n/m) \cdot s$$

  ("complexity" m is proportional to the "information content" s)
- Phase transitions for successful recovery were observed and precisely located by Donoho & Tanner and Stojnic

Observed Phase Transitions

Let $\boldsymbol{x}_0$ be s-sparse, $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_0$ for random $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ $(s < m < n)$.

$$\text{minimize} \quad \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

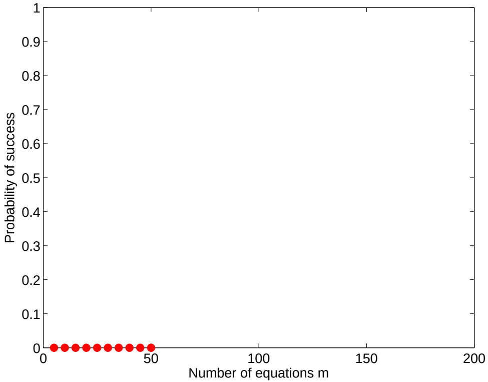


$s = 50,\ m = 25,\ n = 200$

Observed Phase Transitions

Let $\boldsymbol{x}_0$ be s-sparse, $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_0$ for random $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ $(s < m < n)$.

$$\text{minimize} \quad \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

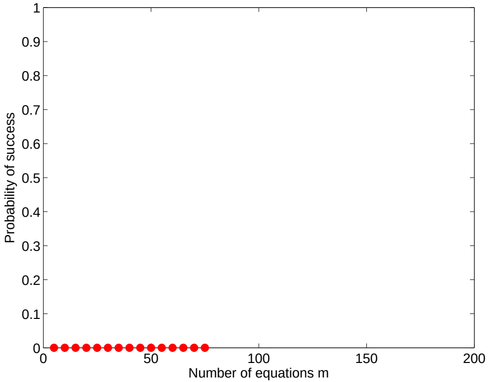


$s = 50,\ m = 50,\ n = 200$

Observed Phase Transitions

Let $\boldsymbol{x}_0$ be s-sparse, $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_0$ for random $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ $(s < m < n)$.

$$\text{minimize} \quad \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

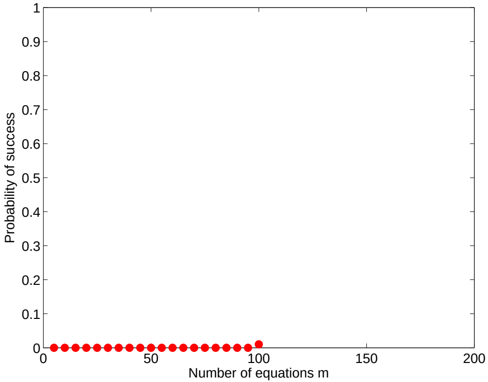


$s = 50,\ m = 75,\ n = 200$

Observed Phase Transitions

Let $\boldsymbol{x}_0$ be s-sparse, $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_0$ for random $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ $(s < m < n)$.

$$\text{minimize} \quad \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

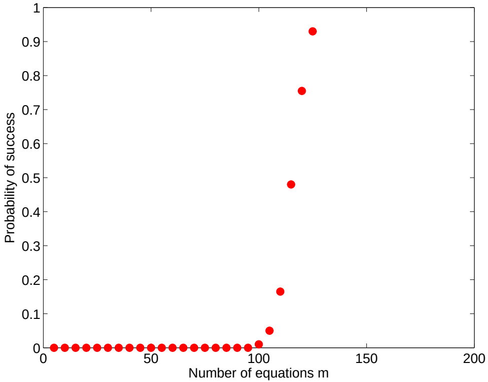


$s = 50,\ m = 100,\ n = 200$

Observed Phase Transitions

Let $\boldsymbol{x}_0$ be s-sparse, $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_0$ for random $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ $(s < m < n)$.

$$\text{minimize} \quad \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

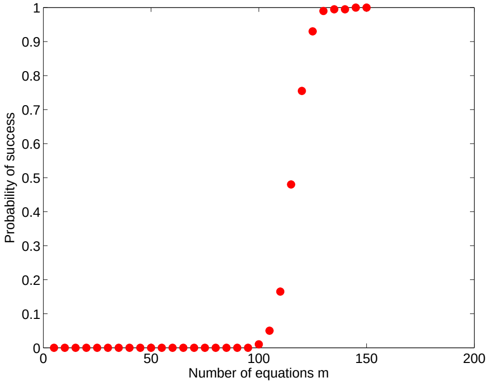


$s = 50,\ m = 125,\ n = 200$

Observed Phase Transitions

Let $\boldsymbol{x}_0$ be s-sparse, $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_0$ for random $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ $(s < m < n)$.

$$\text{minimize} \quad \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

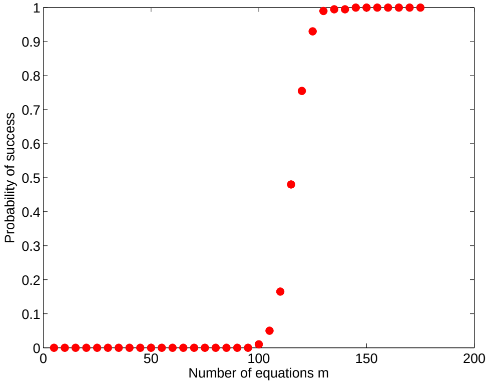


$s = 50,\ m = 150,\ n = 200$

Observed Phase Transitions

Let $\boldsymbol{x}_0$ be s-sparse, $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_0$ for random $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ $(s < m < n)$.

$$\text{minimize} \quad \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

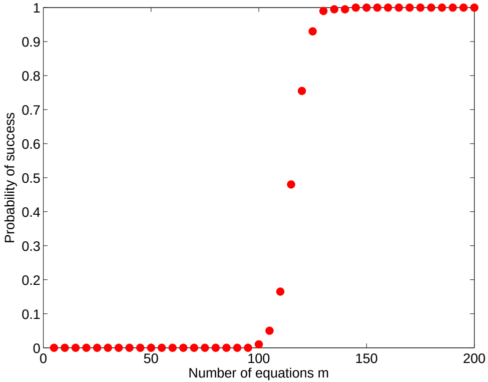


$s = 50,\ m = 175,\ n = 200$

Observed Phase Transitions

Let $\boldsymbol{x}_0$ be s-sparse, $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_0$ for random $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ $(s < m < n)$.

$$\text{minimize} \quad \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$



$s = 50,\ m = 200,\ n = 200$

Phase Transitions for Linear Inverse Problems

Associate to a solution $\boldsymbol{x}_0$ of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ and a convex problem

$$\text{minimize} \quad f(\boldsymbol{x}) \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \qquad (\star)$$

a parameter $\delta(f, \boldsymbol{x}_0)$, the statistical dimension of f at $\boldsymbol{x}_0$.

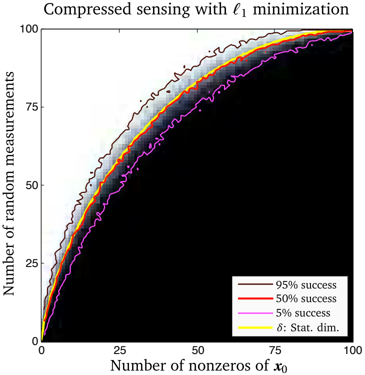
Theorem [Amelunxen, L, McCoy & Tropp, 2014]

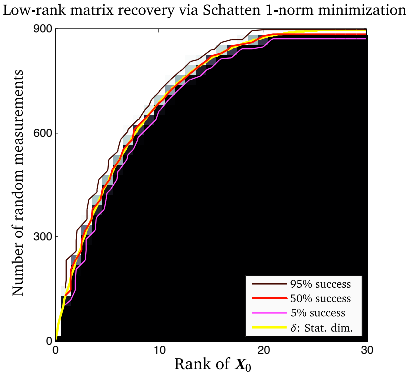
Let $\eta \in (0,1)$ and let $\boldsymbol{x}_0 \in \mathbb{R}^n$ be a fixed vector. Suppose $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ has independent standard normal entries, and let $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_0$. Then

$$m \geq \delta(f, \boldsymbol{x}_0) + a_\eta \sqrt{n} \quad \Longrightarrow \quad (\star) \text{ recovers } \boldsymbol{x}_0 \text{ with probability } \geq 1 - \eta;$$
$$m \leq \delta(f, \boldsymbol{x}_0) - a_\eta \sqrt{n} \quad \Longrightarrow \quad (\star) \text{ recovers } \boldsymbol{x}_0 \text{ with probability } \leq \eta.$$

where $a_\eta := 4\sqrt{\log(4/\eta)}$.

Phase Transitions for Linear Inverse Problems





From Optimization to Geometry

The problem

$$\text{minimize} \quad f(\boldsymbol{x}) \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

has $\boldsymbol{x}_0$ as unique solution if and only if the optimality condition

$$\ker \boldsymbol{A} \cap \mathcal{D}(f, \boldsymbol{x}_0) = \{0\}$$

is satisfied, where

$$\mathcal{D}(f, \boldsymbol{x}_0) := \bigcup_{\tau > 0} \left\{ \boldsymbol{y} \in \mathbb{R}^n \colon f(\boldsymbol{x}_0 + \tau \boldsymbol{y}) \le f(\boldsymbol{x}_0) \right\}$$

is the convex descent cone of f at $\boldsymbol{x}_0$.

- $\boldsymbol{A}$ Gaussian $\Rightarrow \ker \boldsymbol{A}$ uniform in Grassmannian

The Mathematical Problem

Given a closed convex cone $C \subset \mathbb{R}^n$ and a random linear subspace L with $\dim L = k$, find:

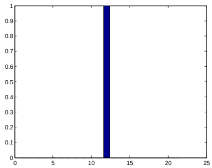
$$\mathbb{P}\{C \cap L \neq \{\mathbf{0}\}\} := \nu_k(\{L \in \mathrm{Gr}(k,n) : C \cap L \neq \{\mathbf{0}\}\}),$$

where ν_k is the normalized Haar measure on the Grassmannian.

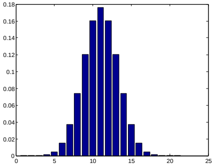
- Bounds can be derived from Gordon's escape through the mesh argument;
- Exact formulas for probability of intersection are based on the Crofton Formula from (spherical) integral geometry.

Spherical Intrinsic Volumes

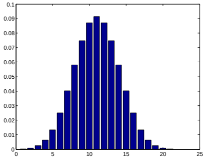
$v_0(C), \dots, v_n(C)$: spherical/conic intrinsic volumes of C.



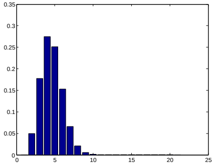
$v_k(L)$, $\dim L = k$



$v_k(\mathbb{R}^n_{\geq 0})$



$v_k(\mathrm{Circ}(n, \pi/4))$



$v_k(\{\boldsymbol{x} \colon x_1 \leq \cdots \leq x_n\})$

The Crofton and Kinematic Formula

- Kinematic Formula

$$\mathbb{P}\{C \cap \boldsymbol{Q}D \neq \{\mathbf{0}\}\} = 2\sum_i v_i(C)\left(\sum_{k \text{ odd}} v_{n-i+k}(D)\right),$$

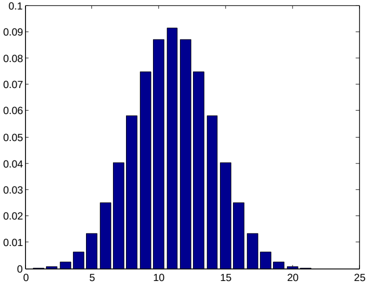
where $\boldsymbol{Q}$ uniformly distributed on $SO(n)$.

- Crofton Formula

$$\mathbb{P}\{C \cap L \neq \{\mathbf{0}\}\} = 2\sum_{k \text{ odd}} v_{m+k}(C),$$

where L uniform in $\mathrm{Gr}(n-m,n)$.

Moments



$\mathrm{Circ}(n, \pi/4)$

Associate to a cone C the discrete random variable X_C with

$$\mathbb{P}\{X_C = k\} = v_k(C),$$

and define the statistical dimension as the expectation

$$\delta(C) = \mathbb{E}[X_C] = \sum_{k=0}^{n} k v_k(C).$$

It appears that X_C might concentrate around $\delta(C)$.

Concentration of Measure

Theorem (Amelunxen-L-McCoy-Tropp 2014)

Let C be a convex cone, and X_C a discrete random variable with distribution $\mathbb{P}\{X_C = k\} = v_k(C)$. Let $\delta(C) = \mathbb{E}[X_C]$. Then

$$\mathbb{P}\{|X_C - \delta(C)| > \lambda\} \leq 4\exp\left(\frac{-\lambda^2/8}{\omega(C)+\lambda}\right) \quad \text{for } \lambda \geq 0,$$

where $\omega(C) := \min\{\delta(C), d - \delta(C)\}$.

- Refined by McCoy-Tropp 2014.
- Central Limit Theorem by Goldstein-Nourdin-Peccati 2017.

Approximate Crofton Formula

Applying the concentration result to the Crofton Formula

Corollary

Let $\eta \in (0,1)$ and assume L uniformly distributed in $\mathrm{Gr}(n-m,n)$. Then

$$\begin{aligned} \delta(C) \leq m - a_\eta \sqrt{n} \quad &\Longrightarrow \quad \mathbb{P}\big\{C \cap L = \{\mathbf{0}\}\big\} \geq 1-\eta; \\ \delta(C) \geq m + a_\eta \sqrt{n} \quad &\Longrightarrow \quad \mathbb{P}\big\{C \cap L = \{\mathbf{0}\}\big\} \leq \eta, \end{aligned}$$

where $a_\eta := 4\sqrt{\log(4/\eta)}$.

- If $C = \mathcal{D}(f, \boldsymbol{x}_0)$, define $\delta(f, \boldsymbol{x}_0) = \delta(C)$.

Approximate Kinematic Formula

Applying the concentration result to the Kinematic Formula

Corollary

Let $\eta \in (0,1)$ and assume one of C, D is not a subspace. Then

$$\begin{aligned}
\delta(C) + \delta(D) \le n - a_\eta \sqrt{n} \quad &\Longrightarrow \quad \mathbb{P}\big\{ C \cap \boldsymbol{Q} D = \{\mathbf{0}\} \big\} \ge 1 - \eta; \\
\delta(C) + \delta(D) \ge n + a_\eta \sqrt{n} \quad &\Longrightarrow \quad \mathbb{P}\big\{ C \cap \boldsymbol{Q} D = \{\mathbf{0}\} \big\} \le \eta,
\end{aligned}$$

where $a_\eta := 4\sqrt{\log(4/\eta)}$.

Euclidean Integral Geometry

Do similar results hold in Euclidean space $\mathbb{R}^n$?

Steiner Formula

$$\mathrm{Vol}_n(K + \lambda \mathsf{B}_n) = \sum_{i=0}^{n} \lambda^{n-i} \kappa_{n-i} \cdot \mathrm{V}_i(K),$$

where K convex body, $\kappa_i = \mathrm{Vol}_i(\mathsf{B}_i)$.

- $\mathrm{V}_i(K)$: i-th intrinsic volume;
- Wills functional $\mathrm{W}(K) = \mathrm{V}_0(K) + \mathrm{V}_1(K) + \cdots + \mathrm{V}_n(K)$.

Crofton and Kinematic Formula

Different normalization (Nijenhuis, ...)

$$\overline{\mathrm{V}}_i(K) := \frac{\omega_{n+1}}{\omega_{i+1}}\,\mathrm{V}_i(K), \qquad \overline{\mathrm{W}}(K) := \sum_{i=0}^{n} \overline{\mathrm{V}}_i(K),$$

where ω_k is the surface area of a k-dimensional unit ball.

- Crofton Formula

$$\int_{\mathrm{Af}(n-i,n)} \overline{\mathrm{W}}(K \cap E)\ \mu_{n-i}(\mathrm{d}E) = \sum_{k=i}^{n} \overline{\mathrm{V}}_k(K).$$

Crofton and Kinematic Formula

Different normalization (Nijenhuis, ...)

$$\overline{\mathrm{V}}_i(K) := \frac{\omega_{n+1}}{\omega_{i+1}}\, \mathrm{V}_i(K), \qquad \overline{\mathrm{W}}(K) := \sum_{i=0}^{n} \overline{\mathrm{V}}_i(K),$$

where ω_k is the surface area of a k-dimensional unit ball.

- Crofton Formula

$$\int_{\mathrm{Af}(n-i,n)} \frac{\overline{\mathrm{W}}(K \cap E)}{\overline{\mathrm{W}}(K)}\; \mu_{n-i}(\mathrm{d}E) = \sum_{k=i}^{n} \frac{\overline{\mathrm{V}}_k(K)}{\overline{\mathrm{W}}(K)}.$$

Crofton and Kinematic Formula

Different normalization (Nijenhuis, ...)

$$\overline{\mathrm{V}}_i(K) := \frac{\omega_{n+1}}{\omega_{i+1}}\,\mathrm{V}_i(K), \qquad \overline{\mathrm{W}}(K) := \sum_{i=0}^{n} \overline{\mathrm{V}}_i(K),$$

where ω_k is the surface area of a k-dimensional unit ball.

- Kinematic Formula

$$\int_{G_n} \overline{\mathrm{W}}(K \cap gM)\; \mu(\,\mathrm{d}g) = \sum_{j=0}^{n} \overline{\mathrm{V}}_j(K) \left(\sum_{k=n-j}^{n} \overline{\mathrm{V}}_k(M) \right).$$

Crofton and Kinematic Formula

Different normalization (Nijenhuis, ...)

$$\overline{\mathrm{V}}_i(K) := \frac{\omega_{n+1}}{\omega_{i+1}}\,\mathrm{V}_i(K), \qquad \overline{\mathrm{W}}(K) := \sum_{i=0}^{n} \overline{\mathrm{V}}_i(K),$$

where ω_k is the surface area of a k-dimensional unit ball.

- Kinematic Formula

$$\int_{G_n} \frac{\overline{\mathrm{W}}(K \cap gM)}{\overline{\mathrm{W}}(K)\,\overline{\mathrm{W}}(M)}\, \mu(\,\mathrm{d}g) = \sum_{j=0}^{n} \frac{\overline{\mathrm{V}}_j(K)}{\overline{\mathrm{W}}(K)} \left(\sum_{k=n-j}^{n} \frac{\overline{\mathrm{V}}_k(M)}{\overline{\mathrm{W}}(K)} \right).$$

Concentration of Intrinsic Volumes

Define the random variable I_K as

$$\mathbb{P}\{I_K = n - i\} = \frac{\mathrm{V}_i(K)}{\mathrm{W}(K)}, \quad \delta(K) := \mathbb{E}[I_K].$$

Theorem (L-Tropp, 2019)

$\mathrm{Var}[I_K] \le 2 \cdot \delta(K)$, and for $t \ge 0$,

$$\mathbb{P}\{|I_K - \delta(K)| \ge t\} \le 2\exp\left(\frac{-t^2/2}{2(\delta(K) + t/3)}\right)$$

- Similar results for $\overline{\mathrm{V}}_i(K)$ and other normalizations.

Approximate Integral Geometry

Theorem (Approximate Crofton Formula)

Let $\eta \in (0,1)$. Then

$$\int_{\mathrm{Af}(n-i,n)} \frac{\overline{\mathrm{W}}(K \cap E)}{\overline{\mathrm{W}}(K)} \; \mu_{n-i}(\mathrm{d}E) \begin{cases} \leq \eta & \bar{\delta}(K) \geq n - i + a_\eta \sqrt{\bar{\delta}(K)} \\ \geq 1 - \eta & \bar{\delta}(K) \leq n - i - a_\eta \sqrt{\bar{\delta}(K)}, \end{cases}$$

where $a_\eta := \sqrt{5 \log(\eta^{-1})}$.

Similar approximate version of Projection Formula.

Approximate Integral Geometry

Approximate Kinematic Formula

Let $\eta \in (0,1)$. Then

$$\int_{G_n} \frac{\overline{\mathrm{W}}(K \cap gM)}{\overline{\mathrm{W}}(K)\,\overline{\mathrm{W}}(M)}\ \mu(\mathrm{d}g) \begin{cases} \leq \eta & \overline{\delta}(K) + \overline{\delta}(M) \geq n + a_\eta \sqrt{\overline{\delta}(K)} \\ \geq 1 - \eta & \overline{\delta}(K) + \overline{\delta}(M) \leq n - a_\eta \sqrt{\overline{\delta}(K)}, \end{cases}$$

where $a_\eta := \sqrt{5 \log(\eta^{-1})}$.

Similar approximate version of Rotation Mean Formula.

Proof of Concentration

Let Y be any random variable.

- Moment generating function (mgf) $m_Y(\theta) := \mathbb{E}\, \mathrm{e}^{\theta Y}$;
- Cumulant generating function (cgf) $\xi_Y(\theta) = \log m_Y(\theta)$.

Cramér-Chernoff

$$\mathbb{P}\{Y \geq t\} \leq \exp\left(\inf_{\theta > 0} [\xi_Y(\theta) - \theta t]\right);$$

$$\mathbb{P}\{Y \leq -t\} \leq \exp\left(\inf_{\theta < 0} [\xi_Y(\theta) + \theta t]\right).$$

- Goal: Find bound on $\xi_Y(\theta)$ when $Y = I_K - \delta(K)$.

Proof Outline

Bound on $\xi_Y(\theta)$ can be deduced from differential inequality:

Proposition

$$\xi''_{I_K}(\theta) \le 2\xi'_{I_K}(\theta) \quad \text{for } \theta \in \mathbb{R}.$$

- Define log-concave probability density

$$\mu_\theta(\boldsymbol{x}) := \frac{1}{\mathrm{W}(K)} \cdot \frac{1}{m_{I_K}(\theta)} \cdot \mathrm{e}^{-V_\theta(\boldsymbol{x})} \quad \text{for } \boldsymbol{x} \in \mathbb{R}^n \text{ and } \theta \in \mathbb{R}.$$

  where $V_\theta(\boldsymbol{x}) := \pi \mathrm{e}^{-2\theta} \operatorname{dist}_K^2(\boldsymbol{x})$.
- Use integral form of Steiner Formula to derive the identities

$$\xi'_{I_K}(\theta) = \mathbb{E}_\theta[2V_\theta], \qquad \xi''_{I_K}(\theta) = \mathrm{Var}_\theta[2V_\theta] - 2\xi'_{I_K}(\theta),$$

  where $\mathbb{E}_\theta$ and Var_θ are mean and variance with respect to μ_θ;

Proof Outline

- Final step: bound the variance term in

$$\xi''_{I_K}(\theta) = \mathrm{Var}_\theta[2V_\theta] - 2\xi'_{I_K}(\theta).$$

- Use Brascamp-Lieb variance inequality

$$\mathrm{Var}_\theta[f] \leq \int_{\mathbb{R}^n} \left\langle (\nabla^2 V_\theta)^{-1} \nabla f,\ \nabla f \right\rangle \mu_\theta(\,\mathrm{d}\boldsymbol{x})$$

to bound

$$\mathrm{Var}[2V_\theta] \leq 4\,\mathbb{E}[2V_\theta] = 4\xi'_{I_K}(\theta).$$

- Concentration of normalized intrinsic volumes along the lines, but uses other (concave) densities and variance bounds (Nguyen, 2014).

For more details:

D. Amelunxen, M. Lotz, M. B. McCoy, and J. A. Tropp
Living on the edge: phase transitions in convex programs with random data.
Information and Inference, 2014

M. B. McCoy and J. A. Tropp
From Steiner formulas for cones to concentration of intrinsic volumes.
Discrete and Computational Geometry, 2014

L. Goldstein, I. Nourdin, and G. Peccati
Gaussian Phase Transitions and Conic Intrinsic Volumes: Steining the Steiner Formula.
The Annals of Applied Probability, 2017

M. Lotz, M. B. McCoy, I. Nourdin, G. Peccati, and J. A. Tropp
Concentration of the Intrinsic Volumes of a Convex Body.
Geometric Aspects of Functional Analaysis - Israel Seminar (GAFA), 2017-2019 (to appear)

M. Lotz and J. A. Tropp
Phase Transitions in Integral Geometry.
In preparation, 2019 (?).

Thank You!