Conditional geometry of the ℓ^p sphere

Steven S. Kim (joint w/ Kavita Ramanan)

http://ssk.im

steven_kim@brown.edu

Division of Applied Mathematics, Brown University, USA

arXiv:1509.05442

BROWN

1. Setup: ℓ^p spheres

Let $p \in [1,\infty]$; $\|\cdot\|_{n,p}$ the ℓ^p norm on $\mathbb{R}^n$; and $\mathbb{S}_{n,p} \doteq \{x \in \mathbb{R}^n : \|x\|_{n,p} = 1\}$. Also:

- sample $X^{(n,p)} \sim$ **cone OR surface measure** on $\mathbb{S}_{n,p}$;
- define $\mu_p(dy) \doteq \frac{1}{2p^{1/p}\Gamma(1+\frac{1}{p})} e^{-|y|^p/p} dy \in \mathcal{P}(\mathbb{R})$.

Prop. (Poincaré-Borel) Fix $k \in \mathbb{N}$. Then, as $n \to \infty$,

$$\mathrm{Law}\left[n^{1/p}(X_1^{(n,p)}, \cdots, X_k^{(n,p)})\right] \Rightarrow \mu_p^{\otimes k}. \qquad \text{(PB)}$$

Q: What if we condition on $\|X^{(n,p)}\|_{n,q}$ small ($q < p$)?

2. Conditional geometry

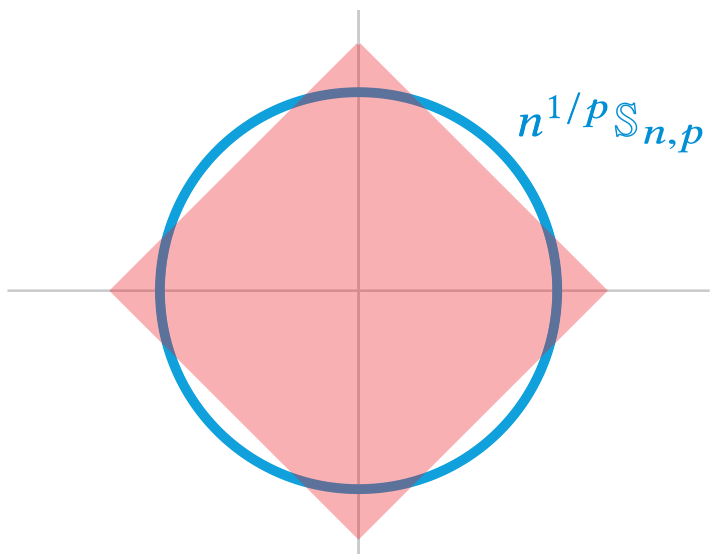
Thm.A (roughly): for $q < p$, in high dimensions, a random point on the ℓ^p sphere conditioned on having small ℓ^q norm is close (in distribution) to a random point drawn from an appropriately scaled ℓ^q sphere.

Thm.A: Let $q < p \in [1,\infty)$, $\beta \le \beta_{p,q} \doteq \frac{1}{q}\left(\Gamma(\frac{1}{q})/\Gamma(\frac{p+1}{q})\right)^{q/p}$. As $n \to \infty$,

$$\mathrm{Law}\left[n^{1/p}(X_1^{(n,p)}, \cdots, X_k^{(n,p)}) \,\middle|\, \left\|n^{1/p-1/q}X^{(n,p)}\right\|_{n,q}^q \le \beta\right]$$
$$\approx \quad \mathrm{Law}\left[\beta^{1/q}\, n^{1/q}(X_1^{(n,q)}, \cdots, X_k^{(n,q)})\right].$$

Note: we mean "$\approx$" in the sense of weak convergence on $\mathcal{P}(\mathbb{R}^k)$.

Key is to analyze probabilities of **"geometric" rare events**.



3. Large deviations for ℓ^p sphere

- empirical measure: $L_{n,p} \doteq \frac{1}{n}\sum_{i=1}^n \delta_{n^{1/p}X_i^{(n,p)}}$.
- relative entropy: $H(\nu\|\mu) \doteq \int_{\mathbb{R}} \log(\frac{d\nu}{d\mu})\, d\nu$ if $\nu \ll \mu$; else, $= +\infty$.

Thm.B: Let $q < p \in [1,\infty]$. The sequence $(L_{n,p})_{n\in\mathbb{N}} \subset \mathcal{P}(\mathbb{R})$ satisfies a *large deviation principle (LDP)* in the Wasserstein-q topology, with convex good rate function (GRF) $\mathbb{H}_p$,

$$\mathbb{H}_p(\nu) \doteq \begin{cases} H(\nu\|\mu_p) + \frac{1}{p}\left[1 - \int_{\mathbb{R}}|x|^p d\nu\right] & \text{if } \int_{\mathbb{R}}|x|^p d\nu \le 1, \\ +\infty & \text{else.} \end{cases}$$

- **Wasserstein-q:** $\gamma_n \Rightarrow \gamma$ (weak) *and* $\int_{\mathbb{R}}|x|^q d\gamma_n \to \int_{\mathbb{R}}|x|^q d\gamma$.
- **LDP:** for all Borel measurable $\Gamma \subset \mathcal{P}(\mathbb{R})$,

$$-\inf_{\nu\in\Gamma^\circ}\mathbb{H}_p(\nu) \le \liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(L_{n,p}\in\Gamma^\circ)$$
$$\le \limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(L_{n,p}\in\bar{\Gamma}) \le -\inf_{\nu\in\bar{\Gamma}}\mathbb{H}_p(\nu).$$

- **GRF:** $\mathbb{H}_p$ has compact level sets.
- cf. **Sanov:** $Y^{(n,p)} \sim \mu_p^{\otimes n}$ (i.i.d.), rate fn. $H(\cdot\|\mu_p)$, weak topology.

4. Pf of Thm.B – cone vs. surface

Lem. (Rachev-Rüschendorf, Schectmann-Zinn) Let $n \in \mathbb{N}$, $p \in [1,\infty]$, and $X^{(n,p)} \sim$ **cone measure** on $\mathbb{S}_{n,p}$. For $Y^{(n,p)} \sim \mu_p^{\otimes n}$,

$$X^{(n,p)} \overset{(d)}{=} \frac{Y^{(n,p)}}{\|Y^{(n,p)}\|_{n,p}}.$$

- *Idea:* Use LDP for $L_{n,p}^Y \doteq \frac{1}{n}\sum_{i=1}^n \delta_{Y_i^{(n,p)}}$ to obtain LDP for $L_{n,p}$.
- *Problem:* can't apply contraction principle to map $L_{n,p}^Y \mapsto L_{n,p}$.
- *Resolution:* Cramér's theorem in $\mathcal{P}(\mathbb{R}) \times \mathbb{R}_+$.

Lem. (Naor-Romik) Let $n \in \mathbb{N}$, $p \in [1,\infty)$. Then,

$$\frac{d\mathrm{surface}_{n,p}}{d\mathrm{cone}_{n,p}}(x) = C_{n,p}\left(\sum_{i=1}^n |x_i|^{2p-2}\right)^{1/2}, \quad x \in \mathbb{S}_{n,p}. \qquad \text{(RN)}$$

Upshot: Can use (RN) and Varadhan's lemma to prove LDP for $(L_{n,p})_{n\in\mathbb{N}}$ under **surface measure** is same as under **cone measure**.

5. Pf of Thm.A – Gibbs conditioning

Gibbs cond. principle (roughly). Assume the LDP of **Thm.B** holds. If we condition on the rare event $\{L_{n,p} \in F\}$ for a suitable closed set $F \subset \mathcal{P}(\mathbb{R})$, then for large n, $L_{n,p}$ concentrates on

$$\mathcal{M}_F \doteq \left\{\nu \in \mathcal{P}(\mathbb{R}) : \mathbb{H}_p(\nu) = \min_{\gamma\in F}\mathbb{H}_p(\gamma)\right\}.$$

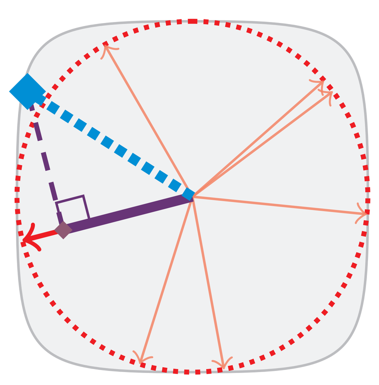
- *Challenge:* which F is "suitable"?
- *Hint:* If $\mathcal{M}_F = \{\mu_q\}$, basic *propagation of chaos* results show that **Thm.B** + **Gibbs cond.** + **Prop. (PB)** $\Rightarrow$ **Thm.A**.
- *Solution:* Pick $F = F_\beta = \left\{\nu \in \mathcal{P}(\mathbb{R}) : \int_{\mathbb{R}}|x|^q d\nu \le \beta\right\}$.

Note 1: We use the fact that **Thm.B** holds w.r.t. Wasserstein-q topology to verify the (unstated) technical conditions on F required for Gibbs conditioning.

Note 2: For "small" $\beta \le \beta_{p,q}$, $\mathcal{M}_{F_\beta} = \{\mu_q\}$. For β very "large", F_β is *not* a rare event, and $\mathcal{M}_{F_\beta} = \{\mu_p\}$. Not so explicit a form for "intermediate" β.

6. App'n of LDP: random projections

Prop. (roughly) A suitably scaled random projection of a uniformly random point from an ℓ^p ball satisfies some large deviation principle with good rate function.



$\xi^{(n,p)} \sim$ uniform on scaled ℓ^p ball $\{x \in \mathbb{R}^n : \|x\|_{n,p} \le n^{1/p}\}$

$\Theta^{(n)} \sim$ uniform on S^{n-1}

"Annealed": $\frac{1}{\sqrt{n}}\langle\xi^{(n,p)}, \Theta^{(n)}\rangle_n$, $n \in \mathbb{N}$, satisfies a LDP with GRF $\mathbb{I}_p^{\mathrm{ann}}$.

"Quenched": Fix non-random $\theta^{(n)} \in S^{n-1}$, $n \in \mathbb{N}$, such that as $n \to \infty$, $\frac{1}{n}\sum_{i=1}^n \delta_{\sqrt{n}\theta_i^{(n)}} \to \nu \in \mathcal{P}(\mathbb{R})$ (w.r.t. Wasserstein-$\frac{p}{p-1}$). Then, $\frac{1}{\sqrt{n}}\langle\xi^{(n,p)}, \theta^{(n)}\rangle_n$, $n \in \mathbb{N}$, satisfies a LDP with GRF $\mathbb{I}_{p,\nu}^{\mathrm{que}}$.

Thm.C (variational formula): For $p > 2$,

$$\mathbb{I}_p^{\mathrm{ann}}(w) = \inf_{\substack{\nu\in\mathcal{P}(\mathbb{R}) \\ \text{w/ density}}}\left\{\mathbb{I}_{p,\nu}^{\mathrm{que}}(w) + \mathbb{H}_2(\nu)\right\}.$$

Proof: Crucially uses LDP for $(L_{n,p})_{n\in\mathbb{N}}$ of **Thm.B**.