$$\text{(conditional) } \mathbb{P} \quad \xleftrightarrow{\text{large dev.}} \quad \text{variational problems}$$

"geometric rare events"

Large Deviations of Random Projections

Steven Soojin Kim

(w/ N. Gantert & K. Ramanan)

Brown University, USA

Classical probability

iid $(X_1, \cdots, X_n) = X^{(n)}$

empirical mean:

$$\frac{1}{n} S_n \doteq \frac{1}{n} \sum_{i=1}^{n} X_i \, .$$

as $n \to \infty$:

(LLN) $\frac{1}{n} S_n \to 0$;

(CLT) $\frac{1}{\sqrt{n}} S_n \Rightarrow N(0,1)$;

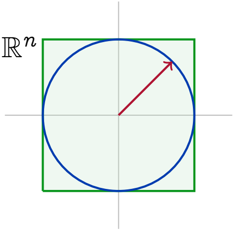
(LDP) $\mathbb{P}(S_n \geq n\, x) \approx \exp(-n\, I(x))$.

Geometric perspective

$X_1, \cdots, X_n \sim$ iid $\mathrm{Unif}([-1,1])$. $X^{(n)} \sim \mathrm{Unif}([-1,1]^n)$ (n-dim. cube).

$$\text{empirical mean} \quad = \textit{projection} \text{ of } X^{(n)} \text{ onto unit vector } \mathbf{1}_n \doteq \frac{1}{\sqrt{n}}(\underbrace{1, 1, \cdots, 1}_{n}) \in S^{n-1}$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{\sqrt{n}}\langle X^{(n)}, \mathbf{1}_n \rangle.$$



Geometric questions

I. General $\theta^{(n)} \in S^{n-1}$? weighted sum $\frac{1}{\sqrt{n}}\langle X^{(n)}, \theta^{(n)}\rangle$.

II. *Dependent* $X^{(n)}$? e.g., uniform on convex body B_n.

KNOWN *CLT*: random proj. of convex body.

$X^{(n)} \sim \mathrm{Unif}(B_n)$ for convex body $B_n \subset \mathbb{R}^n$.

$\Theta^{(n)} \sim \mathrm{Unif}(S^{n-1})$.

For large n, $\langle X^{(n)}, \Theta^{(n)}\rangle \approx$ Gaussian (!!)

[Diaconis + Freedman, Klartag, Bobkov, Meckes, ...]

KNOWN *LDP*: $\mathbf{1}_n$ proj. of product measure.

For $X^{(n)}$ iid $\Rightarrow$ Cramér's thm, $\mathbb{P}(\frac{1}{n}S_n \geq x) \approx \exp(-n\,\Lambda^*(x))$.

Large deviations principles (LDP)

HEURISTIC:

$\mathbb{P}(\frac{1}{n}S_n \geq x) \approx e^{-nI(x)}$

DEF:

$\mathcal{X}$ topological space, $(\xi_n)_{n\in\mathbb{N}}$ seq of $\mathcal{X}$-valued r.v. $\sim$ LDP with *rate function* $I : \mathcal{X} \to [0, +\infty]$ if $I(\cdot)$ l.s.c. &

$$\begin{aligned}\inf_{x\in A^\circ} I(x) &\leq \liminf_{n\to\infty} \tfrac{1}{n} \log \mathbb{P}(\xi_n \in A) \\ &\leq \limsup_{n\to\infty} \tfrac{1}{n} \log \mathbb{P}(\xi_n \in A) \leq \inf_{x\in\overline{A}} I(x).\end{aligned}$$

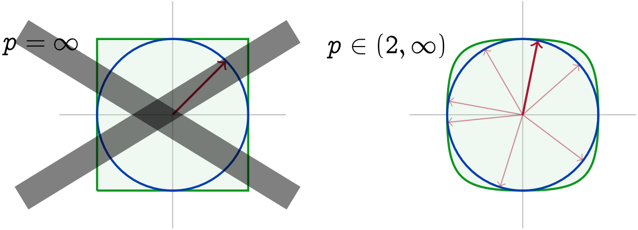
$I(\cdot)$ is a *good* rate function (GRF) if compact level sets.

Why LDPs?

- asymptotic decay of “rare event" prob.
- MANTRA: *“any large deviation is done in the least unlikely of all the unlikely ways!"*

Setup: ℓ^p balls

$B_{n,p} \doteq \{x \in \mathbb{R}^n : \|x\|_{\ell^p} \leq 1\}$;
i.e., $X^{(n,p)}$ uniform on ℓ^p unit ball of $\mathbb{R}^n$.



Question: LDP for projections of ℓ^p balls?

LDP for $\kappa_n \langle X^{(n,p)}, \Theta^{(n)} \rangle$ for $X^{(n,p)} \sim \mathrm{Unif}(\ell^p \text{ ball of } \mathbb{R}^n)$?

(scaling) natural scaling κ_n?

("annealed") LDP for *random* $\Theta^{(n)} \sim \mathrm{Unif}(S^{n-1})$?

("quenched") LDP for *fixed* $\theta^{(n)} \in S^{n-1}$?

- i.e., conditioning on $\{\Theta^{(n)} = \theta^{(n)}, n \in \mathbb{N}\}$.
- how does rate function vary with $\theta^{(n)}$?

(relationship) How do "quenched" and "annealed" relate?

The scaling heuristic

For LDP of $(\xi_n)_{n\in\mathbb{N}}$, want $\xi_n \sim 1/n$ ...

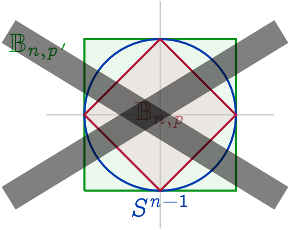
e.g., $\mathrm{Var}(\frac{1}{n}S_n) \sim \frac{1}{n^2}\left(\sum_{i=1}^n 1\right) = \frac{1}{n}$

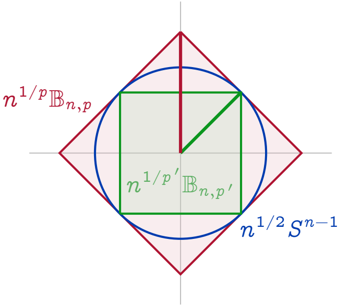
$$\kappa_n\langle X^{(n,p)}, \Theta^{(n)}\rangle$$

Note $X_i^{(n,p)} \sim n^{-1/p}$ and $\Theta_i^{(n)} \sim n^{-1/2}$ ...

so $\mathrm{Var}\langle n^{1/p}X^{(n,p)}, n^{1/2}\Theta^{(n)}\rangle \sim \sum_{i=1}^n (1\cdot 1) = n$ ...

$\Rightarrow$ pick $\kappa_n = \frac{n^{1/2}n^{1/p}}{n}$.





Annealed and Quenched

$$W^{(n,p)} \doteq \kappa_n \langle X^{(n,p)}, \Theta^{(n)} \rangle \quad \text{(random direction)}$$

$$W_\theta^{(n,p)} \doteq \kappa_n \langle X^{(n,p)}, \theta^{(n)} \rangle \quad \text{(condition on fixed direction)}$$

Prop.

(an) $2 \le p \le \infty$: $(W^{(n,p)})_{n\in\mathbb{N}} \sim$ LDP w/ GRF $\mathbb{I}_p^{\mathsf{A}}$.

(qu) $1 < p \le \infty$: for "a.e." directions $(\theta^{(n)})_{n\in\mathbb{N}}$, $(W_\theta^{(n,p)})_{n\in\mathbb{N}} \sim$ LDP w/ GRF $\mathbb{I}_p^{\mathsf{Q}}$.

(for the experts)

- the rate functions are quasiconvex;
- annealed $1 \le p < 2$? *slower speed* of LDP;
- quenched $p = 1$? *domain* problems;
- Cramér directions $(1, 1, \cdots)$ *NOT* in the "a.e." set!!

What does “a.e." mean?

Let $\theta = (\theta^{(n)})_{n\in\mathbb{N}} \in \prod_{n\in\mathbb{N}} \mathbb{R}^n \doteq \mathbb{A}$.

(sequence of directions, or triangular array)

Empirical measure $L_{n,\theta} \doteq \frac{1}{n}\sum_{i=1}^n \delta_{\sqrt{n}\theta_i^{(n)}}$.

Let $\gamma \in \mathcal{P}(\mathbb{A})$. If, for γ-a.e. $\theta \in \mathbb{A}$,

$$L_{n,\theta} \xrightarrow{n\to\infty} \nu \in \mathcal{P}(\mathbb{R}),$$

then, the quenched LDP holds:

$$\gamma\left(\theta \in \mathbb{A} : (W_\theta^{(n,p)})_{n\in\mathbb{N}} \sim \text{LDP w/ GRF } \mathbb{I}_{p,\nu}^{\mathsf{Q}}\right) = 1.$$

Example

σ_n uniform meas. on S^{n-1} ... and $\gamma = \bigotimes_{n\in\mathbb{N}} \sigma_n \in \mathcal{P}(\mathbb{A})$

... then $\nu \sim N(0,1)$ (Poincaré-Maxwell-Borel)

The variational formula

For $\nu \in \mathcal{P}(\mathbb{R})$, $\mu \sim N(0,1)$, let

$$\mathbb{H}(\nu) \doteq \overbrace{H(\nu\|\mu)}^{\text{relative entropy}} + \frac{1}{2}\Big(1 - \overbrace{\int_{\mathbb{R}} x^2 \nu(dx)}^{\text{2nd moment}}\Big), \quad \text{if } \int_{\mathbb{R}} x^2 \nu(dx) \leq 1,$$

$$\doteq +\infty, \quad \text{else.}$$

Lem.

$\mathbb{H}$ is LDP good rate function for $(L_{n,\theta})_{n\in\mathbb{N}}$
(the empirical measure for uniform on S^{n-1})

Thm. For $p > 2$,

$$\mathbb{I}_p^{\mathsf{A}}(w) = \inf_{\nu\in\mathcal{P}(\mathbb{R})} \Big\{ \underbrace{\mathbb{H}(\nu)}_{\text{environment}} + \underbrace{\mathbb{I}_{p,\nu}^{\mathsf{Q}}(w)}_{\text{quenched}} \Big\}.$$

For $p = 2$, $\mathbb{I}_p^{\mathsf{A}} = \mathbb{I}_{p,\mu}^{\mathsf{Q}}$.

The minimizer

Lemma

A unique minimizer exists! For $w \in \mathbb{R}$, let

$$\nu_w \doteq \arg \min_{\nu \in \mathcal{P}(\mathbb{R})} \{ \mathbb{H}(\nu) + \mathbb{I}^{\mathsf{Q}}_{p,\nu}(w) \}.$$

IDEA:

Minimizer ν_w related to conditional law of *environment* Θ, given $W^{(n,p)} \geq w$.

CONJ.

For $p > 2$, as $|w| \to 1$ ... ν_w " $\to$ " $\frac{1}{2}(\delta_{-1} + \delta_{+1})$ ...

i.e., conditional on a very large projection,

$\mathrm{Law}(\theta^{(n)}) \approx \mathrm{Unif}\left(\frac{1}{\sqrt{n}} \{-1, +1\}^n \right)$.

Some concluding thoughts

THE TOOLBOX:

- geometry of ℓ^p balls
- classical large deviations
- (simple) Glivenko-Cantelli and Sanov-type theorems for dependent triangular arrays
- $+ \ldots$

GOAL: *"conditional geometry"*, generally;

(e.g., [asymptotic] behavior of surface measure on ℓ^p sphere, conditioned on small ℓ^q moment?)

Thank you!