Large deviations of random projections

Steven S. Kim, Brown University
(joint with Nina Gantert & Kavita Ramanan)

BROWN

Motivation: geometric view

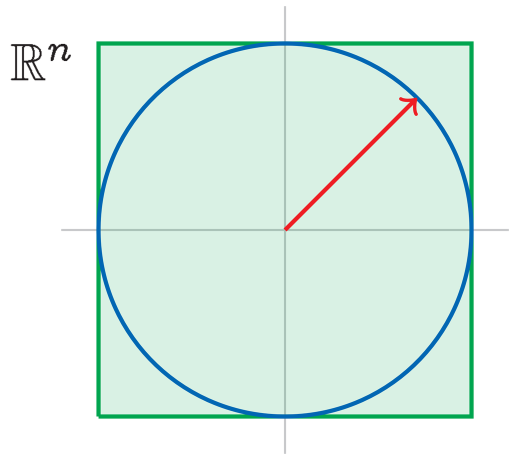
Let $X^{(n)} = (X_1, \cdots, X_n)$ iid random vector.

The empirical mean of (X_i)
$=$
projection of $X^{(n)}$ onto
$\mathbf{1}_n \doteq \frac{1}{\sqrt{n}}(1, \cdots, 1) \in \mathbb{S}^{n-1}$;

i.e.,

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

e.g., $X^{(n)} \sim \mathrm{Unif}([-1,1]^n)$.



A lot known for projections of high-dimensional product measures onto $\mathbf{1}_n$ – e.g., LLN, CLT, LDP.

QUESTIONS:

I. General $\theta^{(n)} \in \mathbb{S}^{n-1}$? i.e., *weighted sums*

$$\tfrac{1}{n}\textstyle\sum_{i=1}^n X_i \sqrt{n}\theta_i^{(n)} = \tfrac{1}{\sqrt{n}}\langle X^{(n)}, \theta^{(n)} \rangle.$$

II. $X^{(n)}$ uniform on other convex body B_n? (in general, this imposes *dependence* of the coordinates $X_1, \cdots, X_n$)

Known

KNOWN: CLT for random proj. of convex body. Let $X^{(n)} \sim \mathrm{Unif}(B_n)$ for convex $B_n \subset \mathbb{R}^n$, and $\Theta^{(n)} \sim \mathrm{Unif}(\mathbb{S}^{n-1})$. Then, for large n, $\langle X^{(n)}, \Theta^{(n)} \rangle \approx$ Gaussian!! [Diaconis + Freedman, Klartag, Bobkov, Meckes, ...]

KNOWN: LDP for $\mathbf{1}_n$ proj. of product measure. For $X^{(n)}$ iid $\Rightarrow$ Cramér's theorem.

Q: LDP for random projection of convex body?

Review: large deviations

MAIN IDEA: $\mathbb{P}(W_n \geq w) \approx \exp(-n\mathbb{I}(w))$.

Let $\mathcal{X}$ a topological space. A sequence of $\mathcal{X}$-valued r.v. $(W_n)_{n\in\mathbb{N}}$ satisfies a *large deviations principle (LDP)* with *rate function* $\mathbb{I}$ if for all measurable A,

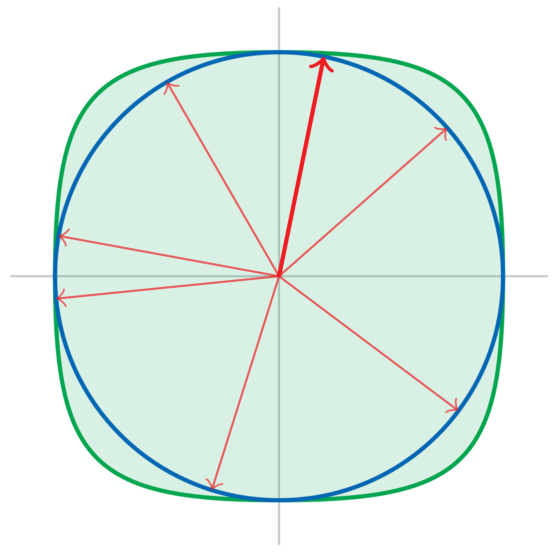
$$-\inf_{w\in A^\circ}\mathbb{I}(w) \leq \liminf_{n\to\infty}\tfrac{1}{n}\log\mathbb{P}(W_n \in A^\circ)$$
$$\leq \limsup_{n\to\infty}\tfrac{1}{n}\log\mathbb{P}(W_n \in \bar{A}) \leq -\inf_{w\in\bar{A}}\mathbb{I}(w)$$

The rate function $\mathbb{I}$ is a *good rate function* if level sets are compact.

Setup: ℓ^p balls

For LDP, need a *sequence* of r.v. e.g.,

$X^{(n,p)}$ uniform on $B_{n,p}$, the ℓ^p unit ball of $\mathbb{R}^n$.
$\Theta^{(n)}$ uniform on $\mathbb{S}^{n-1}$.



(NOTE: for $p \neq \infty$ and large n, law of $X^{(n,p)}$ is a high-dimensional *non-product* measure)

Aside: a spherical Sanov's thm

Let $\alpha^{(n)} = (\alpha_1^{(n)}, \cdots \alpha_n^{(n)})$ uniform on $\sqrt{n}\mathbb{S}^{n-1}$, and $L_n^\alpha \doteq \frac{1}{n}\sum_{i=1}^n \delta_{\alpha_i^{(n)}}$ the empirical measure.

LEM. (Poincaré) $L_n^\alpha \Rightarrow \mu_2 \sim N(0,1)$.

$$\mathbb{H}(\nu) \doteq \begin{cases} \overbrace{\int_{\mathbb{R}} \log\left(\frac{d\nu}{d\mu_2}\right) d\nu}^{\text{relative entropy}} + \frac{1}{2}[1 - \overbrace{\int_{\mathbb{R}} x^2 \nu(dx)}^{\text{2nd moment}}], & \text{if } \nu \ll \mu_2 \text{ and } \int x^2\nu(dx) \leq 1\,; \\ +\infty, & \text{else.} \end{cases}$$

LEM. Let $r < 2$. Then, $(L_n^\alpha)_{n\in\mathbb{N}}$ satisfies an LDP in $\mathcal{P}(\mathbb{R})$ equipped with the Wasserstein-r topology, with the good rate function $\mathbb{H}$.

(NOTE: Similar LDP holds for *cone measure* on $n^{1/q}\partial B_{n,q}$)

Results: the LDP

$$\boxed{\text{Q:}}\ \text{LDP for } (W_{n,p})_{n\in\mathbb{N}}, \text{ where } W_{n,p} \doteq \kappa_{n,p}\langle X^{(n,p)}, \Theta^{(n)} \rangle?$$

THEOREM. Let $p \in [2,\infty)$. Let $\kappa_{n,p} \doteq \frac{n^{1/p}n^{1/2}}{n}$. (NOTE: this scaling so that $W_{n,p}$ is of "order $\frac{1}{n}$")

("annealed") $(W_{n,p})_{n\in\mathbb{N}}$ satisfies an LDP with a quasiconvex good rate function $\mathbb{I}_p^{\mathrm{an}}$.

("quenched") For a.e. $\theta^{(n)} \in \mathbb{S}^{n-1}$, conditioning on *fixed* directions $\Theta^{(n)} = \theta^{(n)}$ for all $n \in \mathbb{N}$, $(W_{n,p})_{n\in\mathbb{N}}$ satisfies an LDP with a quasiconvex good rate function $\mathbb{I}_{p,\mu_2}^{\mathrm{qu}}$.

(relationship) The annealed and quenched rate functions are related through a variational formula.

(NOTE: the a.e. quenched rate function $\mathbb{I}_p^{\mathrm{qu}}$ *does not* depend on the particular $\theta^{(n)}$!)

Variational formula

PROP. For $p > 2$, $\mathbb{I}_p^{\mathrm{an}}(w) =$

$$\inf_{\substack{\nu\in\mathcal{P}(\mathbb{R}) \\ \text{w/ density}}} \{\mathbb{I}_{p,\nu}^{\mathrm{qu}}(w) + \underbrace{\int_{\mathbb{R}} \log\left(\tfrac{d\nu}{dx}\right) d\nu}_{\text{-entropy}}\} + \log\sqrt{2\pi e}\,.$$

For $p = 2$, $\mathbb{I}_p^{\mathrm{an}} = \mathbb{I}_{p,\mu_2}^{\mathrm{qu}}$.

Proof ingredients: let $R_n \doteq \frac{1}{n}\sum_{i=1}^n (\alpha_i^{(n)} Y_i, |Y_i|^p)$.

1. Let $\bar{\Phi}_p$ be the Gärtner-Ellis "pressure functional" for (R_n).,
2. For $r < 2$, compute LDP for $L_n^\alpha \doteq \frac{1}{n}\sum_{i=1}^n \delta_{\alpha_i^{(n)}}$ in Wasserstein-r topology.
3. LDP of 2. + Varadhan's lem to relate $\Psi_{p,\nu}$ and $\bar{\Phi}_p$ via a variational formula.
4. Show $\bar{\Phi}_p^*$ is good rate function for (R_n) by a convexity argument. □

(NOTE: Both $p = 2$ equality and step 4. of proof rely on rotation invariance: i.e., $\langle \Theta^{(n)}, u \rangle \stackrel{(d)}{=} \langle \Theta^{(n)}, v \rangle$ for all $u, v \in \mathbb{S}^{n-1}$.)

Representations of ℓ^p ball

Let $\mu_p(dy) \propto \exp(-|y|^p/p)\,dy$.

- Let $Y^{(n)} = (Y_1, \cdots, Y_n) \stackrel{\text{iid}}{\sim} \mu_p$, and U an independent Uniform$[0,1]$. Then,
$$X^{(n,p)} \stackrel{(d)}{=} U^{1/n}\frac{Y^{(n)}}{\|Y^{(n)}\|_p}.$$
- Let $Z^{(n)} = (Z_1, \cdots, Z_n) \stackrel{\text{iid}}{\sim} N(0,1)$. Then,
$$\Theta^{(n)} \stackrel{(d)}{=} \frac{Z^{(n)}}{\|Z^{(n)}\|_2}.$$

[Schectmann, Zinn '90 and Rachev, Rusendorf '91]

LEM. Mult. by $U^{1/n}$ does not affect LDP; i.e., LDP for $(W_{n,p})$ reduces to LDP for :

$$\frac{n^{1/p}n^{1/2}}{n}\left\langle \frac{Y^{(n)}}{\|Y^{(n)}\|_p}, \frac{Z^{(n)}}{\|Z^{(n)}\|_2} \right\rangle.$$

Annealed + quenched

Let $p \geq 2$. Define certain analogs of log moment generating functions: for $t_0 < \frac{1}{2}$, $t_1 \in \mathbb{R}$, $t_2 < \frac{1}{p}$,

$$\Phi_p(t_0,t_1,t_2) \doteq \log\mathbb{E}_{Y\sim\mu_p, Z\sim\mu_2}[e^{t_0 Z^2 + t_1 ZY + t_2|Y|^p}];$$
$$\Psi_{p,\nu}(t_1,t_2) \doteq \int_{\mathbb{R}} \log\mathbb{E}_{Y\sim\mu_p}[e^{t_1 zY + t_2|Y|^p}]\,\nu(dz).$$

(NOTE: $p \geq 2$ condition used for bounds on log-mgf of $Y \sim \mu_p$)

Let $\Psi_{p,\nu}^*$ and Φ_p^* be their Legendre transforms, and

$$\mathbb{I}_p^{\mathrm{an}}(w) \doteq \inf\left\{\Phi_p^*(\tau_0,\tau_1,\tau_2) : w = \tau_0^{-1/2}\tau_1\tau_2^{-1/p}\right\};$$
$$\mathbb{I}_{p,\nu}^{\mathrm{qu}}(w) \doteq \inf\left\{\Psi_{p,\nu}^*(\tau_1,\tau_2) : w = \tau_1\tau_2^{-1/p}\right\}.$$

PROP. Let $p \in [2,\infty)$: annealed rate function is $\mathbb{I}_p^{\mathrm{an}}$; quenched rate function $\mathbb{I}_{p,\mu_2}^{\mathrm{qu}}$.

Ongoing work

1. the case $p \in [1,2)$;
2. moderate deviations;
3. other sequences.