Math 254A Lecture 6 Notes

Daniel Raban

April 9, 2021

1 Proving the (S2) Condition Via Compact Exhaustion

1.1 Compact exhaustion of convex open sets

Our setting is a σ-finite measure space (M, λ) with a measureable map $\varphi : M \to X$, where $\mathcal{U}$ is the collection of open convex subsets of X. We are trying to measure

$$\lambda^{\times n} \left\{ p \in M^n : \frac{1}{n} \sum_{i=1}^n \varphi(p_i) \in U \right\} = \exp(n \cdot s(U) + o(n))$$

(if it is finite for each each n, otherwise we get $s(U) = \infty$ and RHS $= \infty$). This exceptional case is not an issue if we can guarantee at most exponential growth, e.g. if $\lambda(M) < \infty$.

We can also define a point function

$$s(x) = \inf_{U \ni x} s(U),$$

and this is upper semicontinuous and concave. The next step needs an extra condition:

Each $U \in \mathcal{U}$ is a countable union of compact convex sets.

Here are examples where we can prove this property.

Example 1.1. $X = \mathbb{R}^d$. Let U be convex and open, and let $F_n = \{x \in U : |x| \leq n, \operatorname{dist}(x, U^c) \geq 1/n\}$. This is a closed subset of U (which is bounded and hence compact), and $U = \bigcup_n F_n$. To show that this is convex, we need to make sure the last condition preserves convexity. Observe that this condition holds iff $B_{1/n}(x) \subseteq U$. If this holds at x and y, then

$$B_{1/n}(tx + (1-t)y) = tB_{1/n}(x) + (1-t)B_{1/n}(y) \subseteq U.$$

Example 1.2. $X = Y^*$, where Y is a Banach space and X has the weak*-topology.

To prove the second example, we need the following:

Lemma 1.1. *For $X = Y^*$, if $U \in \mathcal{U}$, then there exist $y_1, \ldots, y_k$ and an open, convex $V \subseteq \mathbb{R}^k$ such that*

$$U = \{x : (\langle x, y_1\rangle, \ldots, \langle x, y_k\rangle) \in V\}$$

i.e. $U = L^{-1}[V]$, where $L : X \to \mathbb{R}^k$ sends $x \mapsto (\langle x, y_1\rangle, \ldots, \langle x, y_k\rangle)$.

Proof. Assume $U \ni 0$, so there exist linearly independent $y_1, \ldots, y_k \in Y$ and a neighborhood W of 0 in $\mathbb{R}^k$ such that $U \supseteq L^{-1}[W]$ (L as above). The main work is showing that $U = L^{-1}[V]$ for some $V \subseteq \mathbb{R}^k$. It is equivalent to show that $U = U + z$ for any $z \in \ker L$.

Suppose $z \in \ker L \subseteq U$ and so $\frac{1}{\varepsilon}z \in U$ for all ε. We have, by convexity, that $U \supseteq (1-\varepsilon)U$ for all $\varepsilon \in [0, 1]$. Similarly, $U \supseteq (1-\varepsilon)U + \varepsilon u$, where $u \in U$. So, in particular,

$$U \supseteq (1-\varepsilon)U + \varepsilon \cdot \frac{1}{\varepsilon}z = (1-\varepsilon)U + z$$

for all ε. Hence,

$$U \supseteq \bigcup_{1>\varepsilon>0} (1-\varepsilon)U + z = U + z.$$

By symmetry, $U = U + z$. $\square$

Proposition 1.1. *$X = Y^*$ has the desired property.*

Proof. Let $U = L^{-1}[V] = \bigcup_n L^{-1}[F_n]$ as above, where $L^{-1}[F_n]$ are weak* closed and convex. By Alaoglu's theorem, $X = \bigcup_n \overline{B_n}$, where $\overline{B_n}$ is compact and convex, and so $U = \bigcup_n (L^{-1}[F_n] \cap \overline{B_n})$. $\square$

1.2 Compact exhaustion implies (S2) condition

Proposition 1.2. *Suppose that X and $\mathcal{U}$ have this property. Then the (S2) condition*

$$s(U) = \sup\{s(K) : K \subseteq U \text{ is compact}\}$$

holds, where

$$s(K) := \inf\{\max_i s(U_i) : U_1, \ldots, U_k \in \mathcal{U}, K \subseteq U_1 \cup \cdots \cup U_k\}.$$

Proof. Recall that if $s(U) > -\infty$, then

$$s(U) = \lim_n \frac{\log \lambda^{\times n}(\{\frac{1}{n}\sum_{i=1}^n \varphi(p_i) \in U\})}{n} = \sup_n \frac{\log \lambda^{\times n}(\{\frac{1}{n}\sum_{i=1}^n \varphi(p_i) \in U\})}{n}.$$

Suppose $a < s(U)$. Then by this latter formulation for $s(U)$, there is some m such that $\log(\lambda^{\times m}(\{\cdots \in U\}))/m > a$. Write $U = \bigcup_k F_k$ where the F_k are compact and convex.

So $\lambda^{\times m}(\{\cdots \in U\}) =\uparrow \lim_k \lambda^{\times m}(\{\cdots \in F_k\})$. So there is a compact convex $F \supseteq U$ with $\frac{\log \lambda^{\times m}(\cdots \in F\})}{m} > a$. By convexity of F, this gives

$$\frac{\log \lambda^{\times \ell m}(\{\frac{1}{\ell m}\sum_{i=1}^{\ell m} \cdots \in F\})}{\ell m} > a$$

for all ℓ. Now suppose $F \subseteq U_1 \cup \cdots \cup U_k$ with the $U_i \in \mathcal{U}$. Then $\lambda^{\times \ell m}(\{\cdots \in F\}) \leq k \max_i \lambda^{\times \ell m}(\{\cdots \in U_i\})$. So

$$\frac{\log \lambda^{\times \ell m}(\{\cdots \in F\})}{\ell m} \leq o(1) + \max_i \underbrace{\frac{\log \lambda^{\ell m}(\{\cdots \in U_i\})}{\ell m}}_{\to s(U_i)}.$$

The lim sup of this as $\ell \to \infty$ is a lower bound on $\max_i s(U_i)$ whenever $F \subseteq U_1 \cup \cdots \cup U_k$. Hence, $s(F) \geq a$. Since a was arbitrary $< s(U)$, we have (S2). $\square$

1.3 Special cases of our construction

Let's take stock of what we have so far: There exists $s : \mathcal{U} \to [-\infty, \infty]$ satisfying (S1) and (S2) such that

$$\lambda^{\times n}\left(p \in M^n : \frac{1}{n}\sum_{i=1}^{n} \varphi(p_i) \in U\right) = \exp(n \cdot s(U) + o(n))$$

as $n \to \infty$ for all $U \in \mathcal{U}$. We also have an upper semicontinuous point function $s : X \to [-\infty, \infty]$ with $s(U) = \sup\{s(x) : x \in U\}$. Also, if $s : \mathcal{U} \to [-\infty, \infty]$ is locally finite, then $s : X \to [-\infty, \infty)$ and is concave.

Here are a few notable special cases:

Example 1.3. Let $M = A$ be a finite alphabet with λ as counting measure. Then $s(U) \leq \log |A|$ for all U, and $\varphi(a) = \delta_a \in P(A)$. Then $\frac{1}{n}\sum_{i=1}^{n} \varphi(a_i)$ is the empirical distribution p_a, and so our conclusion is

$$|T_n(U)| = \exp(n \sup_{p \in U} s(p) + o(n)).$$

Example 1.4. Let $X = \mathbb{R}^d$, and let $\xi_1, \xi_2, \dots$ be iid random variables with values in $\mathbb{R}^d$. So in the background, there is a probability space (M, λ) and measurable $\varphi : M \to \mathbb{R}^d$ such that $(\xi_1, \xi_2, \dots) \overset{d}{=} (\varphi(p_1), \varphi(p_2), \dots)$, where $(p_1, p_2, \dots) \sim \lambda^{\times \infty}$. Then there exists a point function $s : \mathbb{R}^d \to [-\infty, 0]$ such that

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} \xi_i \in U\right) = \exp\left(n \cdot \sup_{x \in U} s(x) + o(n)\right).$$

(Note that $s(x) \leq 0$ for all x because $s(U) \leq \log \lambda(M) = 0$ for all U.) If this event is unlikely (prob $\to 0$ as $n \to \infty$), then the event is called a **large deviation**, and this is the beginning of "Large Deviations Theory."