Math 245B Lecture 19 Notes

Daniel Raban

February 25, 2019

1 Applications of Riesz-Representation to Probability Measures

1.1 Probability measures on compact metric spaces

Let (X, ρ) be a compact metric space. Since $M(X, \mathbb{R}) = C(X, \mathbb{R})^*$, we may write $\mu(f) := \int f \, d\mu$.

Definition 1.1. A **probability measure** on a measurable space $(X, \mathcal{M})$ us a positive measure μ such that $\mu(X) = 1$.

On (X, ρ), let $P(x)$ be the collection of probability measures. Then $P(X) \subseteq M(X, \mathbb{R})$.

Lemma 1.1. *$P(x) = \{\mu \in C(X, \mathbb{R})^* : \|\mu\| \leq 1, \mu(\mathbb{1}_X) = 1\}$, where $\mathbb{1}_X$ is the constant 1 function.*

Proof. $\mu \in P(X)$ iff $\mu \geq 0$ and $\mu(X) = 1$. All the work is in showing $\supseteq$. We just need to show that if $\|\mu\| \leq 1$ and $\mu(X) = 1$, then $\mu \geq 1$. Let $\mu = \mu^+ - \mu^-$ be the Jordan decomposition of μ. Then $\mu(X) = \mu^+(X) - \mu^{-1}(X) = 1$, and $\|\mu\| = |\mu|(X) = \mu^+(X) + \mu^-(X) \leq 1$. So $\mu(X) = \mu^+(X) = 1$, and $\mu^- = 0$. $\square$

Corollary 1.1. *$P(X)$ us compact and metrizable in the weak* topology.*

Proof. $P(X) = B^* \cap \{\mu \in C(X, \mathbb{R})^* : \mu(\mathbb{1}_X) = 1\}$. The latter set is closed, so $P(X)$ is a weak* closed subset of B^*. So Alaoglu's theorem gives us that $P(X)$ is compact. $\square$

Remark 1.1. $(B^*, \mathcal{T}_{\text{weak}^*})$ is metrizable, so $C(X, \mathbb{R})$ is separable (this was a previous application of Stone-Weierstrass).

Remark 1.2. Here are explicit examples of suitable metrics. Le (f_n) be dense in the unit ball of $C(X, \mathbb{R})$. Then

$$\tilde{\rho}(\mu, \nu) = \max_n \left\{ 2^{-n} \left| \int f_n \, d\mu - \int f_n \, d\nu \right| \right\}$$

is a metric. The same works if we replace the max by a sum.

Remark 1.3. Embed $X \to P(X)$ by sending $x \mapsto \delta_x$. This is a homeomorphic embedding with this topology. The key point is that $x_n \to x$ iff $f(x_n) \to f(x)$ for all $f \in C(x)$; that is, $\int f\, d\delta_{x_n} \to \int f\, d\delta_x$.

Theorem 1.1 (Kiylov-Bogoliubor)**.** *Let $X \neq \varnothing$ be a compact metric space, and let $T : X \to X$ be continuous. Then there exists $\mu \in P(X)$ such that $\mu(T^{-1}[A]) = \mu(A)$ for all $A \in \mathcal{B}_X$.*

Proof. Pick $x \in X$. For $n \in \mathbb{N}$, define $\mu_n := n^{-1} \sum_{i=0}^{n-1} \delta_{T^i(x)}$. Now let μ be $\lim_{k \to \infty} \mu_{n_k}$ for some subsequence that converges. Then for all $f \in C(X)$, we have

$$\int f\, d\mu = \lim_k \int f\, d\mu_{n_k} = \lim_k \frac{1}{n_k} \sum_{i=0}^{n_k - 1} f(T^i(x)).$$

Similarly,

$$\int f \circ T\, d\mu = \lim_k \frac{1}{n_k} \sum_{i=0}^{n_k - 1} f(T^{i+1}(x)) = \lim_k \frac{1}{n_k} \sum_{i=1}^{n_k} f(T^i(x)).$$

So

$$\left| \int f\, d\mu - \int f \circ T\, d\mu \right| = \lim_k \frac{1}{n_k} |f(x) - f(T^{n_k}(x))| \leq \frac{2\|f\|_u}{n_k} \to 0.$$

So we get $\int f\, d\mu = \int f \circ T\, d\mu$ for all $f \in C(X)$. This implies that μ is T-invariant (exercise in regularity). $\square$

Remark 1.4. We could write the last step as $\int f\, d\mu = \int f\, d(T_* \mu)$ for all $f \in C(X)$, where $T_* \mu$ is the push-forward measure of μ by T. This gives, $\mu = T_* \mu$.

1.2 Probability measures on non-compact metric spaces

What if our metric space is not compact? One nice way to do things is to work in locally compact spaces. Another important case is to look at complete and separable metric spaces. In either case, it is no longer true that $M(X, \mathbb{R}) = C(X, \mathbb{R})^*$.

Definition 1.2. Let (X, ρ) be a locally compact metric space with $(\mu_n) \subseteq p(X)$ and $\mu \in P(X)$. Then $\mu_n \to \mu$ **vaguely** if $\int f\, d\mu_n \to \int f\, d\mu$ for all $f \in C_0(X, R)$ (functions which tend to 0 at ∞). The **vague topology** is the corresponding topology.

Remark 1.5. The vague topology has a nice Banach space interpretation, but $P(X)$ is usually no longer compact. We can see this by looking at the embedding of $X \to P(X)$.

Remark 1.6. See Folland Proposition 7.19 for an interpretation of the vague topology of $P(\mathbb{R})$ in terms of $F(x) = \mu((-\infty, x])$.

Now suppose X is complete and separable.

Lemma 1.2. *If $\mu \in P(X)$, then for all $\varepsilon > 0$, ther eis a compact $K \subseteq X$ such that $\mu(K^c) < \varepsilon$.*

This motivates the following definition.

Definition 1.3. $\mathcal{A} \subseteq P(X)$ is **tight** if for all $\varepsilon > 0$, there exists a compact $K \subseteq X$ such that $\mu(K^c) < \infty$ for all $\mu \in \mathcal{A}$.

Theorem 1.2 (Prohorov)**.** *Let (μ_n) be a sequence in $P(X)$, and assume that $\{\mu_n : n \in \mathbb{N}\}$ is tight. Then there is a subsequence $(\mu_{n_k})_k$ and a measure $\mu \in P(X)$ such that*

$$\int f \, d\mu_{n_k} \xrightarrow{k\to\infty} \int f \, d\mu$$

for all $f \in BC(X)$.

Remark 1.7. Probabilists usually call the topology related to this type of convergence the **weak topology**.[1] You can instead call this the **BC-topology**.

[1]Please try not to do this.