Lecture 14, Wed, 4/27/2022

Today ⊙ Narrow convergence of probability measures
⊙ Weak lower semicont. of $E_K[\cdot]$

Given: X, Y: Polish spaces, $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$.
$c: X \times Y \to [0, \infty]$: Borel measurable.

Def: $\mathcal{A}(\mu,\nu) = \{\gamma \in \mathcal{P}(X\times Y) : \pi^X_\# \gamma = \mu,\ \pi^Y_\# \gamma = \nu\}$.

$$E_K[\gamma] = \int_{X\times Y} c(x,y)\, d\gamma(x,y).$$

<u>Theorem</u> Assume $c: X\times Y \to [0,\infty]$ is lower semicontinuous. Then $\exists\ \hat{\gamma} \in \mathcal{A}(\mu,\nu)$ s.t. $E_K[\hat{\gamma}] = \min_{\gamma \in \mathcal{A}(\mu,\nu)} E_K[\gamma]$.

<u>Review of signed measures, narrow convergence of probability measures, etc.</u>

Let (Z, d) be a Polish space (i.e., a complete and separable metric space).

<u>Finite, signed Borel measures</u>

$\mu : \mathcal{B}(Z) \to \mathbb{R}$ (not $\mathbb{R} \cup \{+\infty, -\infty\}$) is a finite signed measure on Z if μ is σ-additive: for any $E = \bigcup_{j=1}^{\infty} E_j$ disjoint, all $E_j \in \mathcal{B}(Z)$, $\sum_j \mu(E_j)$ converges absolutely, and $\mu(E) = \sum_j \mu(E_j)$, and $\mu(\phi) = 0$.

Denote

$\mathcal{M}(Z) = \{$all <u>finite</u> signed measures on $Z\}$,

$\mathcal{M}_+(Z) = \{\mu \in \mathcal{M}(Z) : \mu \geq 0\}$,

$\mathcal{P}(Z) = \{\mu \in \mathcal{M}_+(Z) : \mu(Z) = 1\}$.

$\mu \in \mathcal{M}_+(Z) \Leftrightarrow \mu$ is a finite, Borel (positive) measure on Z.

$\mu \in \mathcal{P}(Z)$: μ is a (Borel) probability measure on Z.
$\mathcal{P}(Z) \subseteq \mathcal{M}_+(Z) \subseteq \mathcal{M}(Z)$.
$\mathcal{M}(Z)$ is a vector space. $\mathcal{M}_+(Z)$ and $\mathcal{P}(Z)$ are convex subsets of $\mathcal{M}(Z)$.
$\mu \in \mathcal{M}(Z)$: $\underbrace{|\mu|}_{\text{total variation of } \mu} = \mu^+ + \mu^-$, $\mu = \mu^+ - \mu^-$.

Proposition $\mathcal{M}(Z)$ is a normed vector space with $\|\mu\| = |\mu|(Z)$ if $\mu \in \mathcal{M}(Z)$.

Narrow convergence: $\mu_k, \mu \in \mathcal{P}(Z)$, $\mu_n \to \mu$ narrowly if
$$\int_Z \varphi \, d\mu_k \to \int_Z \varphi \, d\mu \quad \forall \varphi \in C_b(Z),$$
where
$C_b(Z) \triangleq \{ f : Z \to \mathbb{R} \mid f \text{ is bounded and continuous} \}$

Let $\varphi \in C_b(Z)$. Define $\|\varphi\|_\infty = \sup_{z \in Z} |\varphi(z)|$. Then $(C_b(Z), \|\cdot\|_\infty)$ is a Banach space. If $\mu \in \mathcal{M}(Z)$, then $l_\mu(\varphi) = \int_Z \varphi \, d\mu$ $(\varphi \in C_b(Z))$ defines $l_\mu : C_b(Z) \to \mathbb{R}$ and $l_\mu \in [C_b(Z)]^*$. $|l_\mu(\varphi)| \le \|\mu\| \, \|\varphi\|_\infty$ $\forall \varphi$.

If Z is locally compact, then by Riesz's Thm. $\mathcal{M}(Z) = [C_0(Z)]^*$ ($\varphi \in C_0(Z)$ means $\varphi : Z \to \mathbb{R}$ is continuous, and $\forall \varepsilon > 0$: $\{ z \in Z : |\varphi(z)| \ge \varepsilon \}$ is compact.)
In particular, in this case, $l_\mu : C_0(Z) \to \mathbb{R}$, defined by $l_\mu(\varphi) = \int_Z \varphi \, d\mu$, satisfies $\|l_\mu\|_{[C_0(Z)]^*} = \|\mu\|_{\mathcal{M}(Z)}$ $(= |\mu|(Z))$

In general, (Z is a Polish space.) $\forall \mu \in \mathcal{M}_+(Z)$, $l_\mu : C_b(Z) \to \mathbb{R}$ satisfies $\|l_\mu\| = \|\mu\|$, since $\|l_\mu\| \le \|\mu\|$

and choosing $\varphi \equiv 1$, we get $l_\mu(\varphi) = \mu(Z) = \|\mu\|$, and $\|\varphi\| = 1$. Hence $\|l_\mu\| = \|\mu\|$. Thus, $\mathcal{M}_+(Z) \subseteq [C_b(Z)]^*$. In particular, $\mathcal{P}(Z) \subseteq$ unit sphere of $[C_b(Z)]^* \subseteq$ closed unit ball of $[C_b(Z)]^*$, which is compact w.r.t. weak-$*$ topology (Banach-Alaoglu Thm). This weak-$*$ topology is defined by the family of seminorms $\{p_\varphi : \varphi \in C_b(Z)\}$, $p_\phi(l) = |l(\phi)|$ $\forall l \in [C_b(Z)]^*$. With the weak-$*$ topology, $[C_b(Z)]^*$ is a locally convex Hausdorff topological vector space. Call this topology on $\mathcal{M}_+(Z) \subseteq [C_b(Z)]^*$ the <u>narrow topology</u>.

Therefore, the narrow convergence of $\mu_k \in \mathcal{P}(Z)$ to $\mu \in \mathcal{P}(Z)$ is weak-$*$ convergence w.r.t. to the weak-$*$ topology of $[C_b(Z)]^*$, the narrow topology. If $\mu_k \in \mathcal{P}(Z)$ and $\mu \in \mathcal{M}_+(Z)$ are such that

$$\int_Z \varphi \, d\mu_k \to \int_Z \varphi \, d\mu \qquad \forall \varphi \in C_b(Z),$$

then, choosing $\varphi \equiv 1$, we get $\mu_k(Z) \to \mu(Z)$. Hence, $\mu(Z) = 1$, and $\mu \in \mathcal{P}(Z)$. If $\exists l \in [C_b(Z)]^*$ such that $\mu_k \to l$ in the narrow topology i.e., $\forall \varphi \in C_b(Z)$, $l(\phi) = \lim_{k \to \infty} \int_Z \varphi \, d\mu_k$. Then, l may not be in $\mathcal{P}(Z)$.

Note: the narrow convergence is not the weak-$*$ convergence w.r.t. $[C_0(Z)]^*$. If Z is locally compact, then $[C_0(Z)]^* = \mathcal{M}(Z)$ (Riesz Thm). Then for $\mu_k \in \mathcal{P}(Z)$, if convergent to some $\mu \in \mathcal{M}(Z)$, then μ may not be in $\mathcal{P}(Z)$.

Example: $Z = \mathbb{R}$. $\mu_k = \delta_k \in \mathcal{P}(\mathbb{R})$ $(k = 1, 2, \cdots)$. $\mu_k \to 0$ weak-* w.r.t. $[C_0(Z)]^*$, i.e., $\forall \varphi \in C_0(Z)$,

$$\int_Z \varphi \, d\mu_k = \varphi(k) \longrightarrow 0.$$

In the case Z is compact, $C_b(Z) = C(Z)$. and $\mathcal{M}(Z) = [C_b(Z)]^*$. Hence $\mathcal{M}(Z)^* = [C_b(Z)]^{**} \supseteq C_b(Z)$. Hence, the narrow convergence is also the weak convergence.

It turns out that the weak-* topology defined by seminorms $\{P_\varphi : \varphi \in C_b(Z)\}$, i.e., the narrow topology, on $\mathcal{P}(Z)$ is metrizable. The metric is the so-called Prokhorov metric, $d_P(\cdot,\cdot)$, defined for $\mu, \nu \in \mathcal{P}(Z)$ by

$$d_P(\mu,\nu) = \inf\{\alpha > 0 : \mu(A) \le \nu(A_\alpha) + \alpha \text{ and } \nu(A) \le \mu(A_\alpha) + \alpha \quad \forall A \in \mathcal{B}(Z)\},$$

where $A_\alpha = \{x \in Z : d(x, A) < \alpha\}$ if $A \ne \phi$, and $\phi_\alpha = 0$ $\forall \alpha > 0$. The metric is also given by the Wasserstein metric, a family of countable members in $C_b(Z)$; See Ambrosio-Gigli-Savaré (2008): §5.1, and the metric β (Lecture (7)).

Theorem If Z is a Polish space, then $\mathcal{P}(Z)$ is a Polish space with respect to the narrow topology. The convergence $\mu_k \to \mu$ $(\mu_k, \mu \in \mathcal{P}(Z))$ is exactly the narrow convergence. QED

Let us collect some results on the narrow convergence of probability measures.

Theorem Let Z be a metric space, $\mu \in \mathcal{P}(Z)$, and $A \in \mathcal{B}(Z)$. Then
$$\mu(A) = \inf\{\mu(U) : U \subseteq Z : \text{open},\ U \supseteq A\},$$
$$\mu(A) = \sup\{\mu(F) : F \subseteq Z : \text{closed},\ F \subseteq A\}.$$

Proof Let $\mathcal{R}$ be the collection of $A \in \mathcal{B}(X)$ such that
$\mu(A) = \inf\{\mu(U) : A \subseteq U : \text{open}\}$ and
$\mu(A) = \sup\{\mu(F) : A \supseteq F : \text{closed}\}$. Note that $\phi \in \mathcal{R}$ and for any $A \in \mathcal{R}$ and any $\varepsilon > 0$, there exist $F \subseteq Z$: closed, $U \subseteq Z$: open, such that $F \subseteq A \subseteq U$ and $\mu(U \setminus F) < \varepsilon$.

We can verify that $\mathcal{R}$ is a σ-algebra, and $\mathcal{R}$ contains all open sets of Z. Hence, $\mathcal{R} = \mathcal{B}(X)$.

QED

Corollary Let Z be a metric space and $\mu, \nu \in \mathcal{P}(Z)$. Then $\mu = \nu \iff \mu(U) = \nu(U)$ for all open sets $U \subseteq Z$ $\iff \mu(F) = \nu(F)$ for all closed sets $F \subseteq Z$. QED

The above theorem and Ulam's lemma (see next lecture) imply the following:

Theorem If Z is a Polish space then any $\mu \in \mathcal{P}(Z)$ is regular, i.e., for any $A \in \mathcal{B}(Z)$,
$$\mu(A) = \inf\{\mu(U) : U \subseteq Z : \text{open},\ U \supseteq A\},$$
$$\mu(A) = \sup\{\mu(K) : K \subseteq Z : \text{compact},\ K \subseteq A\}.$$

The following is a useful result about convergence of probability measures on metric spaces:

Theorem Let Z be a metric space and all μ_k, $\mu \in \mathcal{P}(Z)$ $(k = 1, 2, \cdots)$. The following are equivalent:

(1) $\mu_k \to \mu$ narrowly;

(2) $\int_Z \varphi \, d\mu_k \to \int_Z \varphi \, d\mu \quad \forall \varphi \in BL(Z)$: bounded Lipschitz functions

(3) $\limsup_{k\to\infty} \mu_k(F) \le \mu(F)$ for all closed $F \subseteq Z$;

(4) $\liminf_{k\to\infty} \mu_k(U) \ge \mu(U)$ for all open $U \subseteq Z$;

(5) $\mu_k(A) \to \mu(A) \quad \forall A \in \mathcal{B}(Z)$ with $\mu(\partial A) = 0$.

__Proof.__ See Dudley's book. __QED__

__Example__ Let $Z = \mathbb{R}$. $\mu_k = \delta_{1/k}$ $(k=1,2,\dots)$, $\mu = \delta_0$, $F = (-\infty, 0]$, and $U = (0,\infty)$. Then, F is closed and U is open.

$$\limsup_{k\to\infty} \mu_k(F) = 0 < \mu(F) = 1.$$

$$\liminf_{k\to\infty} \mu_k(U) = 1 > \mu(U) = 0.$$

__Example__ $Z = [0,1]$, Euclid metric. $\mu_k = \frac{1}{k}\sum_{i=1}^{k} \delta_{i/k} \in \mathcal{P}(Z)$.

If $\varphi \in C_b(Z) = C([0,1])$ then $\int_Z \varphi \, d\mu_k = \frac{1}{k}\sum_{i=1}^{k} \varphi(\frac{i}{k}) \longrightarrow \int_0^1 \varphi \, dm$.

Thus, $\{\mu_k\}$ converges to the Lebegue measure m on $[0,1]$ (a probability measure) narrowly. But, if A is the set of all irrational numbers in $[0,1]$, then $m(A) = 1$. But, $\mu_k(A) = 0 \ \forall k$. Thus, $\mu_k(A) \not\to m(A)$.

We now prove the weak-$*$ (i.e., narrow) lower semicontinuity of the K-cost functional, a result used in the proof of the existence theorem.

__Theorem__ If $c: X \times Y \to [0,\infty]$ is lower semicontinuous, then $E_K : \mathcal{P}(X\times Y) \to [0,\infty]$ is lower semicontinuous w.r.t. the narrow topology of $\mathcal{P}(X\times Y)$.

__Proof__ This follows from a lemma in last lecture. Here, we give a direct proof.

We construct $c_k \in C_b(X\times Y)$:

$$c_k(x,y) = \inf_{x'\in X,\, y'\in Y} \{ c(x',y') \wedge k + k d_X(x,x') + k d_Y(y,y') \}.$$

We have

- $0 \le c_k \le c_{k+1} \le c \wedge k = \min(c,k)$.
- $c_k \in \mathrm{Lip}_b(X\times Y)$, as inf of a family of Lip functions with a uniform Lip const. (with fixed k).
- $c_k \uparrow c$ (pointwise increasing, converging to c).

For the last point: Fix $(x,y)\in X\times Y$, assume W.L.O.G. $A(x,y) := \sup_k c_k(x,y) < \infty$. Now, $\forall k$, let (x_k', y_k') be s.t.

$$c(x_k', y_k') \wedge k + k d_X(x,x_k') + k d_Y(y,y_k') \le c_k(x,y) + \frac{1}{k}.$$

Hence, $k d_X(x,x_k') \le A(x,y) + \frac{1}{k}$. So $d_X(x,x_k') \to 0$.

Similarly, $d_Y(y,y_k') \to 0$. Also,

$$c(x_k', y_k') \wedge k \le c_k(x,y) + \frac{1}{k}.$$

By the lower semicont. of c, and the fact that $c_k \uparrow$,

$$c(x,y) \le \liminf_{k\to\infty} c(x_k', y_k') = \liminf_{k\to\infty} c(x_k', y_k') \wedge k$$
$$\le \liminf_{k\to\infty} \left(c_k(x,y) + \frac{1}{k}\right) = \lim_{k\to\infty} c_k(x,y) \le c(x,y).$$

Now, suppose $\gamma_j \in \Pi(\mu,\nu)$ $(j=1,2,\dots)$, $\gamma \in \mathcal{P}(X\times Y)$ satisfy $\gamma_j \to \gamma$ narrowly in $\mathcal{P}(X\times Y)$. $\forall k$.

$$\liminf_{j\to\infty} E_K[\gamma_j] = \liminf_{j\to\infty} \int_{X\times Y} c(x,y)\, d\gamma_j(x,y)$$
$$\ge \liminf_{j\to\infty} \int_{X\times Y} c_k(x,y)\, d\gamma_j(x,y) \overset{\gamma_j \to \gamma}{=} \int_{X\times Y} c_k\, d\gamma.$$

Finally, by the monotone convergence thm,

$$\liminf_{j\to\infty} E_K[\gamma_j] \ge \sup_k \int_{X\times Y} c_k\, d\gamma = E_K[\gamma].$$

QED