Lecture 14, Wed, 4/27/2022

Today

1. Narrow convergence of probability measures
2. Weak lower semicont. of $E_k[r]

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

Def: $A(\mu, \nu) = \{r \in \mathcal{P}(X \times Y) : \pi_1^X r = \mu, \pi_2^Y r = \nu \}$.
$E_k[r] = \int_{X \times Y} c(x, y) \, d\gamma(x, y)$.

Theorem
Assume $C : \times \times Y \to [0, \infty]$ is lower semi-continuous. Then $\exists \gamma \in A(\mu, \nu)$ s.t. $E_k[\gamma] = \min_{r \in A(\mu, \nu)} E_k[r]$.

Review of signed measures, narrow convergence of probability measures, etc.

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

Finite, signed Borel measures

$\mu : \mathcal{B}(Z) \to \mathbb{R}$ (not $\mathbb{R} \cup \{-\infty, +\infty\}$) is a finite signed measure on $Z$ if $\mu$ is $\sigma$-additive: for any $\bigcup_j E_j$ disjoint, all $E_j \in \mathcal{B}(Z)$, $\sum \mu(E_j)$ converges absolutely, and $\mu(\emptyset) = \sum_j \mu(E_j)$ and $\mu(\emptyset) = 0$.

Denote

$M(Z) = \{ \text{all finite signed measures on } Z \}$,
$M_+(Z) = \{ \mu \in M(Z) : \mu \geq 0 \}$,
$\mathcal{P}(Z) = \{ \mu \in M_+(Z) : \mu(Z) = 1 \}$.

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

\( \mu \in \mathcal{P}(Z) \): \( \mu \) 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 \( \mathcal{M}(Z) \).
\( \mu \in \mathcal{M}(Z) \): \( |\mu| = \mu^+ + \mu^- \), \( \mu = \mu^+ - \mu^- \)
*total variation of \( \mu \).

**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
\[
\int_{Z} \varphi \, d\mu_n \to \int_{Z} \varphi \, d\mu \quad \forall \varphi \in C_b(Z),
\]
where
\[
C_b(Z) = \{ f : Z \to \mathbb{R} \mid f \text{ is bounded and continuous} \}.
\]
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
\( \lambda_{\mu}(\varphi) = \int_{Z} \varphi \, d\mu \) \( (\varphi \in C_b(Z)) \) defines \( \lambda_{\mu} : C_b(Z) \to \mathbb{R} \) and \( \lambda_{\mu} \in (C_b(Z))^* \)
with \( ||\lambda_{\mu}||_{*} = ||\mu||_{\infty} \) for \( \forall \varphi \).

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

In general, if \( Z \) is a Polish space, \( \forall \mu \in \mathcal{M}_+(Z) \),
\( \lambda_{\mu} : C_b(Z) \to \mathbb{R} \) satisﬁes \( ||\lambda_{\mu}||_{C_b(Z)} = ||\mu||_{\mathcal{M}(Z)} \) since \( ||\mu||_{\mathcal{M}(Z)} = ||\mu||_{\mathcal{M}_+(Z)} \).
and choosing \( \varphi = 1 \), we get \( E_0(\varphi) = \mu(\varphi) = \| \varphi \| \), and \( \| \varphi \| = 1 \). Hence \( \| E_0 \| = \| \varphi \| \). Thus, \( E_0(\varphi) \in [C_0(Z)]^* \). In particular, \( \mathcal{P}(Z) \) is the unit sphere of \([C_0(Z)]^*\), which is compact w.r.t. weak-* topology (Banach-Alaoglu Thm). This weak-* topology is defined by the family of seminorms \{ \rho_\varphi : \varphi \in [C_0(Z)]^* \}, \( \rho_\varphi(\varphi) = \| \varphi(\cdot) \| \) \( \forall \varphi \in [C_0(Z)]^* \). With the weak-* topology, \([C_0(Z)]^*\) is a locally convex Hausdorff topological vector space. Call this topology on \( \mathcal{N}_r(\varphi) \subseteq [C_0(Z)]^* \) the narrow topology.

Therefore, the narrow convergence of \( \mu_k : \mathcal{P}(Z) \) to \( \mu : \mathcal{P}(Z) \) is weak-* convergence w.r.t. to the weak-* topology of \([C_0(Z)]^*\), the narrow topology. If \( \mu_k \in \mathcal{P}(Z) \) and \( \mu \in \mathcal{M}_{\mathcal{P}}(Z) \) are such that

\[
\int_Z \varphi \, d\mu_k \rightarrow \int_Z \varphi \, d\mu \quad \forall \varphi \in C_0(Z),
\]

then, choosing \( \varphi = 1 \), we get \( \mu_k(Z) \rightarrow \mu(Z) \). Hence, \( \mu(Z) = 1 \), and \( \mu \in \mathcal{P}(Z) \). If \( \exists \lambda \in \{ \theta \in [C_0(Z)]^* \} \) such that \( \mu_k \rightarrow \lambda \) in the narrow topology i.e., \( \lambda \in [C_0(Z)]^* \), then \( \lambda(\varphi) = \lim_{k \to \infty} \int_Z \varphi \, d\mu_k \). Then, \( \lambda \) 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 \( \mu \) may not be in \( \mathcal{P}(Z) \).
Example. \( Z = \mathbb{R} \), \( \mathcal{M}_k = \mathcal{D}_k \mathcal{P}(\mathbb{R}) (k = 1, 2, \ldots) \). \( \mathcal{M}_k \to 0 \) weak-* u.r.t. \( [\mathcal{C}_0(Z)]^* \), i.e., \( \forall \varphi \in \mathcal{C}_0(Z), \int_Z \varphi d\mu_k = \varphi(k) \to 0 \).

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

It turns out that the weak-* topology defined by seminorms \( \{ \varphi \_\varphi \in \mathcal{C}_b(Z) \} \), 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

\[
d_p(\mu, \nu) = \inf \{ \varepsilon > 0 : \mu(A) \leq \nu(A_{\varepsilon}) + \varepsilon \text{ and } \nu(A) \leq \mu(A_{\varepsilon}) + \varepsilon \text{ for all } A \in \mathcal{B}(Z) \},
\]

where \( A_{\varepsilon} = \{ x \in Z : d(x, A) < \varepsilon \} \) if \( A \neq \emptyset \), and \( \emptyset_{\varepsilon} = \emptyset \) for \( \varepsilon > 0 \).

The metric is also given by the Wasserstein metric, a family of countable members \( \mathcal{C}_b(Z) \). See Ambrosio-Gigli-Savare (2008), §5.1, and the metric \( \beta \) (Lecture 1).

**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. \( \square \)

Let us collect some results on the narrow convergence of probability measures.
Theorem Let $Z$ be a metric space, $μ ∈ Π(Z)$, and $A ∈ Ω(Z)$. Then $μ(A) = \inf \{ μ(U) : U ⊆ Z : \text{open}, U ≥ A \}$.

$μ(A) = \sup \{ μ(F) : F ⊆ Z : \text{closed}, F ≤ A \}$.

Proof Let $P$ be the collection of $A ∈ Ω(Z)$ such that

$μ(A) = \inf \{ μ(U) : A ≤ U : \text{open} \}$ and

$μ(A) = \sup \{ μ(F) : A ≥ F : \text{closed} \}$. Note that $P ≡ P$ and for any $A ∈ P$ and any $ε > 0$, there exist $F ⊆ Z : \text{closed}, U ⊆ Z : \text{open}$ such that $P ≤ A$ and $μ(U \setminus F) < 0$.

We can verify that $P$ is a $σ$-algebra, and $P$ contains all open sets of $Z$. Hence, $P = Ω(Z)$. Q.E.D.

Corollary Let $Z$ be a metric space and $μ, ν ∈ Π(Z)$. Then $μ = ν ↔ μ(U) = ν(U)$ for all open sets $U ⊆ Z$ $↔ μ(F) = ν(F)$ for all closed sets $F ⊆ Z$. Q.E.D.

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

Theorem If $Z$ is a Polish space then any $μ ∈ Π(Z)$ is regular, i.e., for any $A ∈ Ω(Z)$,

$μ(A) = \inf \{ μ(U) : U ⊆ Z : \text{open}, U ≥ A \}$,

$μ(A) = \sup \{ μ(K) : K ⊆ Z : \text{compact}, K ≤ A \}$.

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

Theorem Let $Z$ be a metric space and all $μ_k, μ ∈ Π(Z)$ ($k = 1, 2, \ldots$). The following are equivalent:
(1) $\mu_k \rightarrow \mu$ narrowly;
(2) $\sum_{k=1}^{\infty} \mu_k = \mu$ for $\mathcal{V}$ bounded Lipschitz functions
(3) $\limsup_{k \to \infty} \mu_k(F) \leq M(F)$ for all closed $F \subseteq Z$;
(4) $\liminf_{k \to \infty} \mu_k(U) \geq M(U)$ for all open $U \subseteq Z$;
(5) $\mu_k(A) \rightarrow \mu(A)$ $\forall A \in \mathcal{B}(Z)$ with $\mu(\partial A) = 0$.

Proof. See Dudley's book. Q.E.D

Example. Let $Z = \mathbb{R}$, $\mu_k = \sum_{i=1}^{\infty} \delta_{x_i}$ ($k=1,2,\ldots$), $\mu = \delta_0$, $F = (-\infty, 0)$, and $U = (0, \infty)$. Then, $F$ is closed and $U$ is open.

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

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

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

If $\varphi \in C_b(Z) = C([0,1])$ then $\sum_{k=1}^{\infty} \varphi \mu_k = \frac{1}{k} \sum_{i=1}^{k} \varphi(\frac{i}{k}) \rightarrow \int_{Z} \varphi \, dm$.

Thus, $\{\mu_k\}$ converges to the Lebesgue measure on $[0,1]$. (a probability measure) narrowly. But, if $A$ is the set of all irrational numbers in $[0,1]$, then $\mu(A) = 1$.

But, $\mu_k(A) = 0$ $\forall k$. Thus, $\mu_k(A) \not\to \mu(A)$.

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

Theorem. If $C : X \times Y \rightarrow [0,\infty]$ is lower semi-continuous, then $E_k : \mathcal{P}(X \times Y) \rightarrow [0,\infty]$ is lower semi-continuous w.r.t. the narrow topology of $\mathcal{P}(X \times Y)$.

Proof. We construct $C_k \in C_b(X \times Y)$ so that $C_k \uparrow C$.

For any $k \in \{0, 1, 2, \ldots\}$, set
$c_k(x, y) = \inf_{x', y' \in \mathcal{Y}} \{ c(x', y') + k d_k(x, x') + k d_y(y', y') \}$.

We have

1. $0 \leq c_k \leq c_{k+1} \leq c \land k = \min(c, k)$.
2. $c_k \in \text{Lip}_k(X \times Y)$, as inf of a family of Lip functions with a uniform Lip const. (with fixed $k$).
3. $c_k \uparrow c$ (pointwise increasing, converging to $c$).

For the last point: Fix $(x, y) \in X \times Y$, assume W.l.o.g. $A(x, y) = \sup_k c_k(x, y) < \infty$. Then, for all $k$, let $(x_k, y_k)$ be s.t.

$C(x_k, y_k) \land k + k d_k(x, x_k') + k d_y(y_k, y_k') \leq C(x, y) + \frac{1}{k}$.

Hence, $k d_k(x, x_k') \leq C(x, y) + \frac{1}{k}$. So $d_k(x, x_k') \to 0$. Similarly, $d_y(y_k, y_k') \to 0$. Also,

$c(x_k', y_k') \land k \leq C(x, y) + \frac{1}{k}$.

By the lower semi-cont. of $c$, and the fact that $c_k \uparrow c$, $c(x, y) \leq \liminf_{k \to \infty} C(x_k, y_k) = \liminf_{k \to \infty} C(x_k', y_k') \land k \leq \liminf_{k \to \infty} (C(x, y) + \frac{1}{k}) = \lim_{k \to \infty} C(x, y) \leq C(x, y)$.

Now, suppose $\gamma_j \in \mathcal{A}(\mathcal{C}_u, \nu)$ ($j \geq 1, 2, \ldots$), $\gamma \in \mathcal{P}(X \times Y)$ satisfy $\gamma_j \to \gamma$ narrowly in $\mathcal{P}(X \times Y)$. Then

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

Finally, by the monotone convergence theorem,

\[ \liminf_{j \to \infty} E_k[\gamma_j] \geq \sup_{\gamma \in \mathcal{P}(X \times Y)} \int_{X \times Y} c(x, y) d\gamma = E_k[\gamma]. \quad Q.E.D. \]