5. Measures as functionals

Let $(E, \|\cdot\|)$ be a normed space over $\mathbb{R}$. A function $\varphi : E \to \mathbb{R}$ is a **functional** if it is linear, that is $\varphi(au+bv) = a\varphi(u)+b\varphi(v)$ for any vectors $u, v \in E$ and any scalars $a, b \in \mathbb{R}$. It is easy to check that if φ is continuous then it is bounded on $B_E = \{v \in E : \|v\| \le 1\}$[1]. Indeed, if we suppose that $v_n \in B_E$ satisfy $|\varphi(v_n)| > n$ then taking $u_n = (1/\sqrt{n}) \cdot v_n$ we have $u_n \to 0$ while $\varphi(u_n) \to \infty$, a contradiction. A norm of a continuous functional φ is defined as

$$\|\varphi\| = \sup_{v \in B_E} |\varphi(v)| < \infty.$$

We come back to separable metrizable spaces X. Every $\mu \in P(X)$ is a function $Bor(X) \to \mathbb{R}$ but we are going to treat such a measure μ as a functional on $E = C_b(X)$; recall that $C_b(X)$, the space of all bounded continuous functions on X, is equipped with the supremum norm. Simply μ defines a functional on $C_b(X)$, denoted by the same letter, acting by integration

$$\mu : C_b(X) \to \mathbb{R}, \quad \mu(g) = \int_X g \, \mathrm{d}\mu,$$

Note that $\|\mu\| = 1$ for if $g : X \to [-1.1]$ then clearly $|\mu(g)| \le 1$ and the norm is attained for the constant function $g = 1$. The space $C_b(X)$ carries an additional structure (is a Banach lattice) which is a partial order: $f \le g$ means that $f(x) \le g(x)$ for every $x \in X$. The functional given by a measure $\mu \in P(X)$ is **positive**, meaning that $\mu(g) \ge 0$ for every $g \in C_b(X)$ such that $g \ge 0$.

We note first that every $\mu \in P(X)$ is determined by the values of the integral on continuous functions.

Lemma 5.1. *If $\mu, \nu \in P(X)$ satisfy $\mu(g) = \nu(g)$ for every $g \in C_b(X)$ then $\mu = \nu$.*

Proof. Consider any open set $V \subseteq X$. We know that $V = \bigcup_n F_n$ for some increasing sequence of closed sets F_n. For every n there is (!!) a continuous function $g_n : X \to [0, 1]$ such that $g_n|F_n = 1$ and $g_n|X \setminus V = 0$. It follows that g_n converge pointwise to the characteristic function χ_V. We have

$$\int_X g_n \, \mathrm{d}\mu \to \mu(V), \quad \int_X g_n \, \mathrm{d}\nu \to \nu(V),$$

by the Lebesgue dominated convergence theorem; hence $\mu(V) = \nu(V)$.

Once we know that μ and ν agree on every open set, we get $\mu(B) = \nu(B)$ for every $B \in Bor(X)$ by (outer-)regularity of measures, see Theorem 4.1. $\square$

[1] here we consider 'closed' balls

Now the question is if we can go the other way round: given a nice functional φ on $C_b(X)$, is there a measure $\mu \in C_b(X)$ such that $\varphi(g) = \int_X g \, \mathrm{d}\mu$ for every $g \in C_b(X)$? This is a subject of **the Riesz representation theorem** which holds for any compact space X (also in the nonmetrizable case). Hopefully, we shall come back to the general version later; for the time being we can note a simple proof in the metrizable case, with a bit of cheating.

Theorem 5.2. *Let K be a compact metrizable space. If $\varphi : C(K) \to \mathbb{R}$ is a norm-one positive functional then there is a unique $\mu \in P(K)$ which represents φ, i.e.*

$$\varphi(g) = \int_K g \, \mathrm{d}\mu \quad \textit{for every } g \in C(K).$$

Proof. Uniqueness follows from Lemma 5.1.

Consider first $K = 2^{\mathbb{N}}$ — the proof in this case is quite simple because we have enough simple continuous functions:

Define $\mu(C) = \varphi(\chi_C)$ for every clopen $C \subseteq 2^{\mathbb{N}}$. Then μ is additive, since for disjoint clopens C_1, C_2 we have $\chi_{C_1 \cup C_2} = \chi_{C_1} + \chi_{C_2}$. Hence μ is a finitely additive probability measure on $\mathrm{clop}(2^{\mathbb{N}})$ and it extends uniquely to a measure on $Bor(2^{\mathbb{N}})$ (denoted still by μ); recall that μ is continuous from above on $\emptyset$ for free. This is the required measure: The formula $\varphi(g) = \int_K g \, \mathrm{d}\mu$ holds for any simple continuous function g by linearity of φ and linearity of the integral. For any $f \in C(2^{\mathbb{N}})$ and $\varepsilon > 0$ there is a simple continuous function g such that $\|f - g\|_\infty < \varepsilon$. Then

$$\int_K |f - g| \, \mathrm{d}\mu \le \varepsilon, \quad \text{and } |\varphi(f) - \varphi(g)| = |\varphi(f - g)| \le \varepsilon,$$

which implies that $\varphi(f)$ and $\int_K f \, \mathrm{d}\mu$ differ by at most 2ε. So, finally, they are equal.

Consider now an arbitrary metrizable compact space K. By a result proved before the virus, there is a continuous surjection $\theta : 2^{\mathbb{N}} \to K$. Consider

$$E = \{g \circ \theta : g \in C(K)\};$$

this is a subspace of $C(2^{\mathbb{N}})$. Given a functional φ on $C(K)$, we can define a functional ψ on E simply by $\psi(g \circ \theta) = \varphi(g)$ for $g \in C(K)$. Such ψ can be extended to a positive norm-one functional ψ' on $C(2^{\mathbb{N}})$. A cheating factor: this is a version of the so called Hahn-Banach theorem, it is not immediate but let us believe it.

Now we are fine: By the first part of the proof, there is $\nu \in P(2^{\mathbb{N}})$ such that $\psi'(f) = \int_{2^{\mathbb{N}}} f \, \mathrm{d}\nu$ for $f \in C(2^{\mathbb{N}})$. Take the image measure $\mu = \theta[\nu]$; then for any $g \in C(K)$ we have

$$\int_K g \, \mathrm{d}\mu = \int_{2^{\mathbb{N}}} g \circ \theta \, \mathrm{d}\nu = \psi'(g \circ \theta) = \psi(g \circ \theta) = \varphi(g).$$

The first equality above follows by the general formula for 'changing the variable in the integral', see List 4. $\square$

Remark 5.3. One can conclude form the Riesz theorem that **every** continuous functional φ on $C(K)$ (K compact) is represented by some finite signed measure, i.e. there are finite measures μ, ν on K such that, for $g \in C(K)$, $\varphi(g)$ equals to the integral of g over $\mu - \nu$.

6. Converging sequences of measures

If $\mu_n, \mu \in P(X)$ we might say that $\mu_n \longrightarrow \mu$ (*pointwise*, or better to say, *setwise*) if $\mu_n(B) \to \mu(B)$ for every $B \in Bor(X)$. This is **not** what we are going to consider here. That type of convergence, considered in functional analysis is of lesser importance.

Definition 6.1. We say that a sequence of measures $\mu_n \in P(X)$ converges to a measure μ and write $\mu_n \longrightarrow \mu$ if μ_n converge as functionals on $C_b(X)$, that is

$$\mu_n(g) = \int_X g \, \mathrm{d}\mu_n \longrightarrow \int_X g \, \mathrm{d}\mu = \mu(g),$$

for every $g \in C_b(X)$.

This is the only type of convergence of measures that we discuss and we should give it a name. The problem is that in probability it is traditionally called a *weak* convergence while in functional analysis it is actually *weak** convergence (while the weak one means something different:-). Let us consider a few examples. recall that δ_x denotes the Dirac measure, a point mass at $\{x\}$.

Example 6.2. For any X and $x_n \in X$, $\delta_{x_n} \longrightarrow \delta_x$ if and only if $\lim_n x_n = x$.

Indeed, simply $\int_X g \, \mathrm{d}\delta_x = g(x)$ and note that $\lim_n x_n = x$ is equivalent to saying that $\lim_n g(x_n) = g(x)$ for every $g \in C_b(X)$ (right?).

Example 6.3. Consider

$$\mu_n = 1/n \sum_{k=1}^{n} \delta_{k/n} \in P([0,1].$$

Then $\mu_n \longrightarrow \lambda$, where λ is the Lebesgue measure on $[0,1]$. Indeed,

$$\int_0^1 g \, \mathrm{d}\mu_n = 1/n \sum_{k=1}^{n} g(k/n) \longrightarrow \int_0^i g \, \mathrm{d}\lambda,$$

because in the middle we have the Riemann sum of a continuous function.

Example 6.4. Let μ_n be a normalized λ restricted to $[0, 1/n]$, i.e.

$$\mu_n(B) = n \cdot \lambda(B \cap [0, 1/n]).$$

Then it is easy (I hope) to check that $\mu_n \longrightarrow \delta_0$.

Recall that $\mu \in P(X)$ is said to be a **continuous measure** if $\mu(\{x\}) = 9$ for every $x \in X$, and μ is **disrete** if μ is concentrated on some countable set (see the list of problems). The above examples show that discrete measures can converge to a continuous measure and vice versa. Note also that $\mu_n \longrightarrow \mu$ does not imply that $\lim_n \mu_n(U) = \mu(U)$ even for open sets U; in fact $\lim_n \mu_n(U)$ may not exist – check that it happens for 6.2.

In the following characterization of convergence, a boundary of a set $B \subseteq X$ is $\overline{B} \setminus \mathrm{int}(B)$.

Theorem 6.5. *Let X be any separable metrizable space and let $\mu_n, \mu \in P(X)$. TFAE*

(i) $\mu_n \longrightarrow \mu$;

(ii) $\limsup_m \mu_n(F) \leq \mu(F)$ *for every closed* $F \subseteq X$*;*
(iii) $\liminf_m \mu_n(V) \geq \mu(V)$ *for every open* $V \subseteq X$*;*
(iv) $\lim_n \mu_n(B) = \mu(B)$ *for every Borel set having* μ*-null boundary.*

Proof. $(i) \to (ii)$. Take a closed set $F \subseteq X$ and $\varepsilon > 0$. There is open $V \supseteq F$ such that $\mu(V \setminus F) < \varepsilon$. Take a continuous $g : X \to [0,1]$ such that $g|F = 1$ and $g|X \setminus V = 0$ (using Tietze's extension theorem, define g on F and $X \setminus V$ as required and exend it continuously). Then for large n

$$\mu_n(F) \leq \int_X g \,\mathrm{d}\mu_n \leq \int_X g \,\mathrm{d}\mu + \varepsilon \leq \mu(V) + \varepsilon \leq \mu(F) + 2\varepsilon.$$

Note that the line above implies (ii),

The equivalence $(ii) \leftrightarrow (iii)$ is a consequence of $\limsup_n \mu_n(F) = 1 - \liminf_n \mu_n(X \setminus F)$.

$(iii) \to (iv)$. Taking $F = \overline{B}$ and $V = \mathrm{int}(B)$ we have $\mu(F \setminus V) = 0$. Hence

$$\liminf_n \mu_n(B) \geq \liminf_n \mu_n(V) \geq \mu(V) = \mu(B) = \mu(F) \geq \limsup_n \mu_n(F),$$

so $\lim_n \mu_n(B) = \mu(B)$.

$(iv) \to (ii)$. Consider a compatible metric ρ on X. Note that for any closed $F \subseteq X$, the boundary of the set

$$B(F,r) = \{x \in X : \rho(x,F) \leq r\}$$

is contained in the 'sphere'

$$S(B,r) = \{x \in X : \rho(x,F) = r\}.$$

Since $S(F,r)$ are pairwise disjoint for different values of r, the set of those r for which $\mu(S(F,r)) > 0$ is countable. If follows that there is a sequence of positive reals $r_k \to 0$ such that $\mu(S(F,r_k)) = 0$. We conclude that

$$\limsup_n \mu_n(F) \leq \lim_n \mu_n(B(F,r_k)) = \mu(B(F,r_k),$$

for every k. But $\lim_k \mu(B(F,r_k) = \mu(F)$, so we are done.

$(ii) \to (i)$. Note first that to prove that $\mu_n \longrightarrow \mu$ it suffices to check the convergence on every continuous $g : X \to [0,1]$.

For such g fix some k and set $F_j = \{x \in X : g(x) \geq j/k\}$. Then

$$1/k \sum_{j=1}^{l} \chi_{F_j} \leq g \leq 1/k \sum_{j=0}^{l} \chi_{F_j},$$

clever, isn't it? We have

$$\limsup_n \int_F g \,\mathrm{d}\mu_n \leq \limsup_n 1/k \sum_{j=0}^{k} \mu_n(F_j) \leq 1/k \sum_{j=0}^{k} \limsup_n \mu_n(F_j) \leq$$

$$\leq 1/k \sum_{j=0}^{k} \mu(F_j) \leq \mu(F_0)/k + \int_X g \,\mathrm{d}\mu,$$

for every k, and this gives $\lim_n \int_X g \,\mathrm{d}\mu_n = \int_X g \,\mathrm{d}\mu$. □