B The Carathéodory extension theorem

For now we leave aside integration, and quickly cover techniques for constructing measures directly. I'm not including proofs in this version of the notes; you can find them in the older notes in §8.

Definition B.1. An *exterior measure* is a function $\mu_* : 2^X \to [0, \infty]$ such that

- $\mu_*(\emptyset) = 0$,
- if $E \subset F$, then $\mu_*(E) \leq \mu_*(F)$, and
- μ_* is 'countably subadditive', i.e.

$$\mu_*\left(\bigcup_i E_i\right) \leq \sum_i \mu_*(E_i).$$

Definition B.2. Given an exterior measure μ_*, we say a set $E \subset X$ is *measurable* if for every subset $A \subset X$,

$$\mu_*(A) = \mu_*(A \cap E) + \mu_*(A \setminus E).$$

(Notice that we always have an inequality.)

Theorem B.3. *The measurable sets for an exterior measure form a σ-algebra, and the restriction of the exterior measure to the measurable sets becomes countably additive (that is, an honest measure).*

In fact, the measures obtained in this way are always complete.

Unfortunately for a general exterior measure it's hard to identify the measurable sets, and there is no guarantee that the Borel sets are all measurable. In the situation that X is a metric space, we can solve this problem with an extra condition.

Definition B.4. An exterior measure μ_* on a metric space X is called a *metric exterior measure* if $\mu_*(A \cup B) = \mu_*(A) + \mu_*(B)$ whenever $d(A, B) > 0$.

Theorem B.5. *All Borel sets are measurable with respect to a metric exterior measure.*

Definition B.6. A *premeasure* on a Boolean algebra $\mathcal{M}_0$ of subsets of X is a function $\mu_0 : \mathcal{M}_0 \to [0, \infty]$ such that

- $\mu_0(\emptyset) = 0$,
- if E_i is a countable family of disjoint sets in $\mathcal{M}_0$, and $\bigcup_i E_i$ also happens to be in $\mathcal{M}_0$, then

$$\mu_0\left(\bigcup_i E_i\right) = \sum_i \mu_0(E_i).$$

Lemma B.7. *From a premeasure we can construct an exterior measure μ_* by*

$$\mu_*(E) = \inf\left\{\sum \mu_0(E_i) \mid E \subset \bigcup_i E_i \text{ and each } E_i \in \mathcal{M}\right\}.$$

Indeed, every set in $\mathcal{M}_0$ is measurable for μ_*, and so by the Carathéodory extension theorem we obtain an extension of μ_0 to a measure μ on the σ-algebra $\mathcal{M}$ generated by $\mathcal{M}_0$.

(This was how you originally constructed Lebesgue measure!)

Next we prove that this extension is unique; as long as μ_0 is σ-finite. (A problem on the 3rd assignment has you construct a counterexample in the case μ_0 is not σ-finite.)

Lemma B.8. *Let μ_0 be a pre-measure on a Boolean algebra $\mathcal{M}_0$, and write $\mu : \mathcal{M} \to [0, \infty]$ for the measure constructed in the previous lemma. Suppose $\mu' : \mathcal{M} \to [0, \infty]$ is some other measure on the σ-algebra generated by $\mathcal{M}_0$.*

Finally suppose the (X, μ_0) is σ-finite. Then $\mu(E) = \mu'(E)$ for any $E \in \mathcal{M}$.

Proof: This is just a classic example of a proof where knowing the standard tricks in analysis can lead you inevitably to the right answer. I'll write a proof here that isn't the 'tidied up' proof you might usually see, but instead an analysis of how to 'follow your nose' through the proof.

To begin with, our job is to prove two real numbers are equal, so often a good idea is to try proving that $\mu(E) \leq \mu'(E)$ and $\mu'(E) \leq \mu(E)$. But which one first? Clearly we should do whatever is easiest first! Recall $\mu(E) = \inf\{\sum \mu_0(E_i) \mid E \subset \bigcup_i E_i\}$, that is, $\mu(E)$ is defined as the infimum of some quantity. There's an obvious strategy for proving an equality of the form $A \leq \inf_x B(x)$ — just prove $A \leq B(x)$ for every x. So let's begin there, showing $\mu'(E) \leq \mu(E)$.

Suppose we have $E \in \mathcal{M}$, and $\{E_i\} \subset \mathcal{M}_0$ such that $E \subset \bigcup_i E_i$. We want to show $\mu'(E) \leq \sum_i \mu_0(E_i)$. It's obvious how to start: since μ' is a measure, it's monotonic, so we have

$$\mu'(E) \leq \mu'\left(\bigcup_i E_i\right)$$

and then by countable additivity

$$\leq \sum_i \mu'(E_i)$$

which, by the hypothesis that μ' is an extension of μ_0, becomes

$$= \sum_i \mu_0(E_i).$$

Now, taking the infimum over all such $\{E_i\}$, we obtain $\mu'(E) \leq \mu(E)$, as desired.

Next, we need to show $\mu'(E) \geq \mu(E)$. How do we prove an inequality of the form $A \geq \inf_x B(x)$? A first attempt might just be to find some x_0 so $A \geq B(x_0)$. Unfortunately, we can smell that ths approach is not going to work here — we know we're secretly aiming to prove an equality (not just the inequality we're locally working on now), and if the infimum isn't actually realised as a minimum (which seems plausible in this setting), this proof strategy would give us something too strong: that $A > \inf_x B(x)$. So we need to be sneakier!

Another thing we can do is give ourselves an epsilon of room, and just prove $A \geq inf_x B(x) - \epsilon$ for every $\epsilon > 0$. Then it suffices to show that $A \geq B(x_0) - \epsilon$ for some x_0, and a very plausible candidate is to take some x_0 so that $B(x_0) \leq \inf_x B(x) + \epsilon$ (i.e. something close to realising the infimum).

We should pause here and worry, however — anytime we want to use an 'epsilon of room' argument, we should check for infinite quantities. Proving the $A \geq \infty - \epsilon$ for every $\epsilon > 0$ is no different than proving $A = \infty$, and so this approach couldn't actually make any progress. So we better show, or assume, that the infimum is actually finite. Hmm... why is that reasonable? In our particular situation, we'd need to be assuming that $\mu(E) < \infty$, i.e. that there is some $\{E_i\} \subset \mathcal{M}_0$ with $E \subset \bigcup E_i$ so that $\sum_i \mu_0(E_i) < \infty$. This actually sounds quite harmless, as the lemma we're trying to prove already gives us a clue via the so-far-unnecessary hypothesis that (X, μ_0) is σ-finite. [1]

In fact, at this point in trying to construct this proof, I'd go off and check that σ-finiteness really does allow me to get away with this assumption — even though in a 'polished' proof I'd probably arrange things as 'First, let us assume that $\mu(E) < \infty$ ... Finally, we treat the case that $\mu(E) = \infty$. ...'

So let's imagine we've proved that $\mu(E) = \mu'(E)$ for every $E \in \mathcal{M}$ with $\mu(E) < \infty$. Now take some $E \in \mathcal{M}$ with $\mu(E) = \infty$. Using σ-finiteness, we can write the whole space X as $X = \bigcup_j X_j$, with each $X_j \in \mathcal{M}_0$ and $\mu_0(X_j) < \infty$. In fact, we can arrange for these X_j to be disjoint, by taking set differences. Now, returning to our E, we can write it as $E = \bigsqcup_j (E \cap X_j)$, and calculate

$$\begin{aligned}\mu(E) &= \mu\left(\bigsqcup_j (E \cap X_j)\right) \\ &= \sum_j \mu(E \cap X_j)\end{aligned}$$

[1] Of course, if you were breaking new ground here, rather than trying to fill in a proof, instead you'd think at this point 'Oh, why don't I just assume σ-finiteness for a while. Maybe later I can work on getting rid of this assumption, if it seems to get in the way of appying this lemma where I need it.'

$$\begin{aligned}&= \sum_j \mu'(E \cap X_j) \\ &= \mu'\left(\bigsqcup_j (E \cap X_j)\right) \\ &= \mu'(E).\end{aligned}$$

Easy!

Okay, now getting back on track we pick $\{E_i\}$ so $E \subset \bigcup_i E_i$ and

$$\sum_i \mu_0(E_i) \leq \mu(E) + \epsilon' < \infty \tag{B.1}$$

and our goal is to prove that $\mu'(E) \geq \sum_i \mu_0(E_i) - \epsilon$. (I gave myself a new ϵ' here, as we can't yet know how it's going to need to be related to the 'target' ϵ. Of course a polished proof would unhelpfully elide this, and just pick ϵ' appropriately when it is introduced.)

Unfortunately at this point it's not obvious what to do. We want to show $\mu'(E)$ is bigger than something, but it's unclear where we could obtain any inequality of that form. The slightly subtle trick is to use the inequality we proved in the first half of this theorem, that $\mu'(F) \leq \mu(F)$ for any $F \in \mathcal{M}$. How could that possibly be useful, given we're after the reverse inequality? Well, we could do something like this, for some set $E \subset H$:

$$\mu'(E) = \mu'(H) - \mu'(H \setminus E) \geq \mu'(H) - \mu(H \setminus E)$$

to at least get started. Considering this, it seems like a great idea, as we've got a good candidate for the larger set H, namely $H = \bigcup_i E_i$, which looks very promising as all the sets E_i and in $\mathcal{M}_0$, so $\mu'(E_i) = \mu(E_i)$. Let's see how that goes:

$$\begin{aligned}\mu'(E) &= \mu'\left(\bigcup_i E_i\right) - \mu'\left(\bigcup_i E_i \setminus E\right) \\ &\geq \mu'\left(\bigcup_i E_i\right) - \mu\left(\bigcup_i E_i \setminus E\right) \\ &= \mu'\left(\bigcup_i E_i\right) - \mu\left(\bigcup_i E_i\right) + \mu(E) \\ &\geq \mu'\left(\bigcup_i E_i\right) - \mu\left(\bigcup_i E_i\right) + \sum_i \mu_0(E_i) - \epsilon' && \text{by Equation (B.1)}\end{aligned}$$

That seems quite promising: $\mu'(\bigcup_i E_i) - \mu(\bigcup_i E_i)$ looks almost like the sort of thing we can control, because each E_i is in $\mathcal{M}_0$. Unfortunately, of course, the infinite union isn't in $\mathcal{M}_0$, so we can't directly relate μ' and μ yet. The obvious trick, however, is to approximate both of these by finite expressions. Of course, we can simply throw out all by finitely many sets in the first term, obtaining

$$\mu'(E) \geq \mu'\left(\bigcup_{i=0}^{N} E_i\right) - \mu\left(\bigcup_i E_i\right) + \sum_i \mu_0(E_i) - \epsilon'$$

for any N we like. As $\mu(\bigcup_i E_i) < \infty$, for sufficiently large N we have

$$\mu\left(\bigcup_{i=0}^{N} E_i\right) \geq \mu\left(\bigcup_i E_i\right) - \epsilon''$$

and thus

$$\mu'(E) \geq \mu'\left(\bigcup_{i=0}^{N} E_i\right) - \mu\left(\bigcup_{i=0}^{N} E_i\right) + \sum_i \mu_0(E_i) - \epsilon' - \epsilon''.$$

Finally choosing ϵ' and ϵ'' so $\epsilon' + \epsilon'' = \epsilon$ gives the desired result. □

C Product measures

Given measure spaces $(X_1, \mathcal{M}_1, \mu_1)$ and $(X_2, \mathcal{M}_2, \mu_2)$ we would like to construct a σ-algebra and measure μ on $X_1 \times X_2$, with the property that $\mu(A \times B) = \mu_1(A)\mu_2(B)$ for every $A \in \mathcal{M}_1$ and $B \in \mathcal{M}_2$.

Said in the language of integrals, given a Radon integral $\int_i$ on X_i, we would like to construct a Radon integral $\int$ on $X_1 \times X_2$, with the property that

$$\int f(x)g(y) = \left(\int_1 f\right)\left(\int_2 g\right)$$

for $f \in C_c(X_1)$ and $g \in C_c(X_2)$.

We now have two very different routes we could taking, depending on whether we wanted to work on the measure-theoretic or integration-theoretic side. As we've spent lots of time already on the 'highbrow' abstract integral viewpoint, let's have a change of scenery, and understand the 'conventional' approach. I like the presentation in Terry Tao's 'An introduction to measure theory' over Stein and Shakarchi's — these notes follow Tao.

Very briefly, however, let me sketch how the product integral is defined. The Stone-Weierstrauss theorem shows that when A and B are compact spaces, the space

$$C_c(A) \otimes C_c(B)$$

consisting of finite sums of functions of the form $f(x)g(y)$, where $f \in C_c(X)$ and $g \in C_c(Y)$, is actually dense (with respect to supremum norm) in $C_c(A \times B)$. When A and B are compact, Radon integrals $\int_A$ and $\int_B$ on them are actually bounded linear functionals. We can define $\int_A \otimes \int_B : C_c(A) \otimes C_c(B) \to \mathbb{R}$ by

$$\left(\int_A \otimes \int_B\right)(f \otimes g) = \left(\int_A f\right)\left(\int_B g\right).$$

The density result shows that there is actually a *unique* bounded linear functional extending this to all of $C_c(A \times B)$, and this extension is our product Radon integral. We finally need to do some extra work to get rid of the assumptions that A and B were compact spaces, but this can be done without too much suffering (indeed, we don't even need to assume our spaces X and Y are σ-compact). You can find all the details of this construction in §6.6 of Pedersen's *Analysis Now*.

Definition C.1. Given two σ-algebras $(X_1, \mathcal{M}_1)$ and $(X_2, \mathcal{M}_2)$ the product σ-algebra $\mathcal{M}_1 \times \mathcal{M}_2$ on $X_1 \times X_2$ is the σ-algebra generated by sets of the form $E_1 \times E_2$, where $E_1 \in \mathcal{M}_1$ and $E_2 \in \mathcal{M}_2$.

Exercise. This is the coarsest σ-algebra so that the projection maps $X_1 \times X_2 \xrightarrow{\pi_i} X_i$ are measurable.

Theorem C.2. *Let $(X_1, \mathcal{M}_1, \mu_1)$ and $(X_2, \mathcal{M}_2, \mu_2)$ by σ-finite measure spaces. Then there is a measure $\mu_1 \times \mu_2$ on the σ-algebra $\mathcal{M}_1 \times \mathcal{M}_2$ satisfying $\mu(A \times B) = \mu_1(A)\mu_2(B)$ for every $A \in \mathcal{M}_1$ and $B \in \mathcal{M}_2$.*

Proof: Our strategy is to define a premeasure, on the Boolean algebra $\mathcal{B}_0$ generated by product sets $A \times B$, for $A \in \mathcal{M}_1$ and $B \in \mathcal{M}_2$, and then to use Lemma B.7 to construct a measure via the Carathéodory extension theorem.

First observe that this Boolean algebra is easy to describe explicitly: it is just the collection of all finite unions of sets $A \times B$ with $A \in \mathcal{M}_1$ and $B \in \mathcal{M}_2$, or indeed, just the collection of all finite disjoint unions of such sets. We define the premeasure then by

$$\mu_0\left(\bigsqcup_{j=0}^{m} E_j \times F_j\right) = \sum_{j=0}^{m} \mu_1(E_j)\mu_2(F_j).$$

One needs to verify that this is actually well defined; that is, for a set S in the Boolean algebra $\mathcal{B}_0$, it doesn't matter how we write it is a disjoint union of measurable rectangles. Once we've done that, it's easy to see μ_0 is finitely additive as a function $\mathcal{B}_0 \to [0, \infty]$.

We need to check μ_0 is a pre-measure, i.e. that is $S \in \mathcal{B}_0$ can be written as a countable disjoint union of $S_i \in \mathcal{B}_0$, then $\mu_0(S) = \sum_i \mu_0(S_n)$. In fact, by writing S as a finite disjoint union of measurable rectangles, and intersecting each S_i with each of these rectangles, we see it suffices to consider the case that S is a measurable rectangle. Moreover, using finite additivity of μ_0 it's enough to handle the case where each S_i is also a measurable rectangle. We thus need to show that if

$$E \times F = \bigsqcup_i E_i \times F_i$$

then

$$\mu_1(E)\mu_2(F) = \sum_i \mu_1(E_i)\mu_2(F_i).$$

At each point (x, y), we have

$$\chi_E(x)\chi_F(y) = \sum_i \chi_{E_i}(x)\chi_{F_i}(y)$$

and we can integrate this over X_2 and apply the monotone convergence theorem to obtain

$$\begin{aligned}\chi_E(x)\mu_2(F) = \int_{X_2} \chi_E(x)\chi_F(y) &= \int_{X_2} \sum_i \chi_{E_i}(x)\chi_{F_i}(y) \\ &= \sum_i \chi_{E_i}(x)\mu_2(F_i).\end{aligned}$$

Now integrating over X_1 and applying the monotone convergence theorem again we have

$$\begin{aligned}\mu_1(E)\mu_2(Y) &= \int_{X_1} \sum_i \chi_{E_i}(x)\mu_2(F_i) \\ &= \sum_i \mu_1(E_i)\mu_2(F_i)\end{aligned}$$

as desired.

We now use Lemma B.7 to extend μ_0 to a measure defined on the product σ-algebra $\mathcal{M}_1 \times \mathcal{M}_2$. □

Corollary C.3. *When $(X_1, \mathcal{M}_1, \mu_1)$ and $(X_2, \mathcal{M}_2, \mu_2)$ are both σ-finite, one sees that the premeasure constructed in the previous proof is also σ-finite, and so by Lemma B.8 there is in fact a unique measure on $X_1 \times X_2$ having the 'correct' value on measurable rectangles.*

C.1 The Fubini-Tonelli theorem

We're next going to prove the Fubini-Tonelli theorem, which we can think of as consisting of the following three statements:

Theorem C.4 (Tonelli). *Suppose $(X_1, \mathcal{M}_1, \mu_1)$ and $(X_2, \mathcal{M}_2, \mu_2)$ are σ-finite measure spaces, and $f : X \times Y \to [0, \infty]$ is a measurable function.*

Then

$$\begin{aligned}\int_{X \times Y} f &= \int_X \left(x \mapsto \int_Y (y \mapsto f(x, y))\right) \\ &= \int_Y \left(y \mapsto \int_X (x \mapsto f(x, y))\right)\end{aligned}$$

(in $[0, \infty]$), and in particular the one-variable functions appearing above are in fact measurable and non-negative, so in each case the second integral makes sense.

The proof of this theorem contains the real work of this section, so we'll put it off for a moment!

Theorem C.5 (Fubini)**.**

Proof: [[...]] □

Theorem C.6 ('Fubini-Tonelli')**.**

Proof: [[...]] □

The key technical tool for our proof of the Tonelli theorem (following Tao's book) is following

Definition C.7. A *monotone class* is a collection of subsets of X which is closed under taking

- countable increasing unions, and
- countable decreasing intersections.

(Obviously a monotone class is also closed under taking finite increasing unions or finite decreasing intersections.) One sees easily enough that any σ-algebra is a monotone class (being closed under taking arbitrary countable unions or countable intersections).

Definition C.8. The monotone class generated by a collection of subsets $\mathcal{X}$, which we'll denote as $\langle \mathcal{X} \rangle_{\text{monotone}}$, is the intersection of all monotone classes containing $\mathcal{X}$ (this is a non-empty intersection because the discrete monotone class 2^X contains any collection $\mathcal{X}$).

The point of introducing monotone classes is the following result, which you should think as as providing an alternative (easier!) description of the σ-algebra generated by a Boolean algebra.

Lemma C.9 (Monotone class lemma). *Suppose $\mathcal{A}$ is a Boolean algebra of subsets of X. Then we have*

$$\langle X\rangle_{monotone} = \langle X\rangle_\sigma.$$

Proof: Unsuprisingly, we prove $\langle\mathcal{A}\rangle_{\text{monotone}} \subseteq \langle\mathcal{A}\rangle_\sigma$ and then $\langle\mathcal{A}\rangle_{\text{monotone}} \supseteq \langle\mathcal{A}\rangle_\sigma$.

It's easy to see that every σ-algebra is also a monotone class.[2] From the definition of $\langle\mathcal{A}\rangle_{\text{monotone}}$ as the intersection of all monotone classes containing $\mathcal{A}$, this gives the first inclusion.

Harder is to show $\langle\mathcal{A}\rangle_{\text{monotone}} \supseteq \langle\mathcal{A}\rangle_\sigma$, but following this pattern we also achieve by showing that $\langle\mathcal{A}\rangle_{\text{monotone}}$ is a σ-algebra.

First we claim $\langle\mathcal{A}\rangle_{\text{monotone}}$ is closed under taking complements. Observe that if take the complement of every set in a monotone class C, we obtain a monotone class which we write as C^c. As $\mathcal{A}$ is closed under taking complements, we have $\mathcal{A} \subset \langle\mathcal{A}\rangle^c_{\text{monotone}}$, and thus $\langle\mathcal{A}\rangle_{\text{monotone}} \subseteq \langle\mathcal{A}\rangle^c_{\text{monotone}}$, establishing the claim.

Second, for each $E \in \mathcal{A}$, define

$$C_E = \{F \in \langle\mathcal{A}\rangle_{\text{monotone}} \mid F \setminus E, E \setminus F, E \cap F, (E \cup F)^c \text{ are all in } \langle\mathcal{A}\rangle_{\text{monotone}}\}.$$

Easily $\mathcal{A} \subset \mathcal{E}_E$ for any E, since $\mathcal{A}$ is a Boolean algebra.

Exercise. Show that C_E is itself a monotone class.

Thus $\langle\mathcal{A}\rangle_{\text{monotone}} \subseteq C_E$ for any $E \in \mathcal{A}$, and so $C_E = \langle\mathcal{A}\rangle_{\text{monotone}}$ for every E. Now define

$$\mathcal{D} = \{E \in \langle\mathcal{A}\rangle_{\text{monotone}} \mid F \setminus E, E \setminus F, E \cap F, (E \cup F)^c \text{ are all in } \langle\mathcal{A}\rangle_{\text{monotone}} \text{ for every } F \in \langle\mathcal{A}\rangle_{\text{monotone}}\}.$$

Since $C_E = \langle\mathcal{A}\rangle_{\text{monotone}}$ for every E, we have $\mathcal{A} \subset \mathcal{D}$.

Exercise. Show that $\mathcal{D}$ is itself a monotone class.

Hence $\mathcal{D} = \langle\mathcal{A}\rangle_{\text{monotone}}$. That is, we've just shown that $\langle\mathcal{A}\rangle_{\text{monotone}}$ is closed under differences, intersections, and 'complements of unions'. As we previously checked it was closed under complements, we have that $\langle\mathcal{A}\rangle_{\text{monotone}}$ is closed under finite unions, too.

Finally, we need to show that $\langle\mathcal{A}\rangle_{\text{monotone}}$ is closed under countable unions.

[[...]] □

[2] Unless you're working in intuitionistic logic, without the law of the excluded middle — which case you ought to just add countable intersections to the definition of a σ-algebra.

Corollary C.10. *If a monotone class contains a Boolean algebra, it also contains the entire σ-algebra generated by that Boolean algebra.*

With this in hand, we can give a rather slick:

Proof of Theorem C.4: By the monotone convergence theorem, and the assumption of σ-finiteness, it suffice to prove the theorem in the case that X_1 and X_2 have finite measure.

By the monotone convergence theorem, it is enough to prove the theorem for simple functions, and thence by linearity enough to prove the theorem for a characteristic function χ_S, for some $S \in \mathcal{M}_1 \times \mathcal{M}_2$.

Now, let C denote the set of all $S \in \mathcal{M}_1 \times \mathcal{M}_2$ such that the theorem holds for χ_S.

By the monotone convergence theorem, C is closed under taking countable increasing unions. Now, since X_1 and X_2 have finite measure, we may also use the downwards monotone convergence theorem, and this shows that C is closed under taking countable decreasing intersections.

Finally, it is easy to see that C contains all the measurable rectangles (by a direct calculation of all the integrals), and indeed that it contains all finite disjoint unions of measurable rectangles. But these sets were exactly the Boolean algebra which generated $\mathcal{M}_1 \times \mathcal{M}_2$ as a σ-algebra (see the second paragraph of the proof of Theorem C.2), so by Corollary C.10, C contains all of $\mathcal{M}_1 \times \mathcal{M}_2$, and we've proved the theorem.

(Pow!) □