CONTENTS

Go over extended reals

1 BASIC DEFINITIONS

Intuitively, σ-algebras contain the subsets of a space which we care about measuring.

Definition 1. Let $P(X)$ denote the power set of X. Then $\mathcal{A} \in P(X)$ is a σ-**algebra** on X if

1. $\varnothing, X \in \mathcal{A}$;
2. if $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$;
3. if $\{A_i\} \subset \mathcal{A}$ is countable then $\bigcup_i A_i \in \mathcal{A}$.

Each $A \in \mathcal{A}$ is a **measurable set**.

Proposition 1. If $\{\mathcal{A}_i\}$ is an arbitrary collection of σ-algebras on X, then so is $\bigcap_i \mathcal{A}_i$.

This lets us define a σ-algebra generated by a set of subsets.

Definition 2. Let $\mathcal{M} \subset P(X)$ be a family of subsets of X. Then $\sigma(\mathcal{M})$ is the σ-algebra **generated by** $\mathcal{M}$, defined by the intersection of all σ-algebras on X containing each element of $\mathcal{M}$.

Example 1. Suppose $(X, \mathcal{T}) \in$ **Top**, then $B(X) := \sigma(\mathcal{T})$ is the **Borel** σ-**algebra**.

Measures are maps that measure an element of a σ-algebra (a measurable subset of a space). We want such maps to have intuitive properties of volume. The main one is that we can calculate the volume of something by breaking it up into (perhaps countably infinite) subvolumes and measuring those volumes instead.

Definition 3. $(X, \mathcal{A})$ is a **measurable space**. A map $\mu : \mathcal{A} \to [0, \infty]$ is a **measure** if

1. $\mu(\varnothing) = 0$;
2. $\sum_{i=0}^{\infty} \mu(A_i) = \mu\left(\cup_{i=1}^{\infty} A_i\right)$ for any countable collection of disjoint A_i (μ is "σ-additive").

$(X, \mathcal{A}, \mu)$ is a **measure space**.

Example 2. The **counting measure** is given by

$$\mu(A) := \begin{cases} \text{number of elements in } A & \text{if } A \text{ is finite,} \\ \infty & \text{else.} \end{cases}$$

Example 3. The **Dirac measure** for $p \in X$ is

$$\delta_p(A) := \begin{cases} 1 & p \in A, \\ 0 & \text{else.} \end{cases}$$

We also want to define the "normal" measure on $\mathbb{R}^n$. It should definitely have the following two properties:

1. $\mu([0,1]^n) = 1$ (the unit hypercube has measure 1);
2. $\mu(x + A) = \mu(A)$ for all $x \in \mathbb{R}^n$ (measure is translation-invariant).

Note by extension, this first property becomes $\mu([a,b]^n) = (b-a)^n$ for all $a < b$. As it turns out, it's impossible to construct such a measure on all of $P(\mathbb{R}^n)$. Instead, we'll have to construct it on a particular σ-algebra inside of $P(\mathbb{R}^n)$. We'll call this the **Lebesgue measure**.

Proposition 2. The Lebesgue measure does not exist on all of $P(\mathbb{R}^n)$, except for the trivial measure $\mu = 0$.

Proof. We'll show something more general: Let $I_0 := (0,1]$, and suppose $\mu(I_0) = \varepsilon < \infty$ and $\mu(x + A) = \mu(A)$. Define equivalence classes on $\mathbb{R}$ by $x \sim y \iff x - y \in \mathbb{Q}$. Note that this forms uncountably many equivalence classes. Define $A \subset I_0$ by choosing one element from each equivalence class (this requires the axiom of choice). We'll now form countably many "shifted" versions of A that cover I_0, and use these to show that I_0 has measure 0 (from this, it'll follow that $\mu(\mathbb{R}) = 0$).

Let $A_i := r_i + A$, where $\{r_i\}$ enumerates $\mathbb{Q} \cap (-1, 1]$, then $I_0 \subset \cup_i A_i \subset (-1, 2]$. Then by σ-additivity and translation invariance,

$$\varepsilon = \mu(I_0) \le \mu\left(\cup_i A_i\right) = \sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty} \mu(A) \le \mu((-1,2]) = 3\varepsilon.$$

But $\sum_{i=1}^{\infty} \mu(A) = \infty$ if $\mu(A) > 0$, so $\mu(A) = 0$. Then $\mu(I_0) \le \mu\left(\cup_i A_i\right) = 0$, and

$$\mu(\mathbb{R}) = \mu\left(\bigcup_{n\in\mathbb{N}} (n + I_0)\right) = \sum_{n\in\mathbb{N}} \mu(n + I_0) = \sum_{n\in\mathbb{N}} 0 = 0.$$

Thus, $\mu = 0$, as any set $X \in P(\mathbb{R})$ has measure $\mu(X) \le \mu(\mathbb{R}) = 0$. **Higher dimensions** $\square$

2 THE LEBESGUE INTEGRAL

Definition 4. For measurable spaces $(\Omega_i, \mathcal{A}_i)$, a map $f : \Omega_1 \to \Omega_2$ is **measurable** if

$$f^{-1}(A_2) \in \mathcal{A}_1 \text{ for all } A_2 \in \mathcal{A}_2.$$

Describe intuition here w/ integral figure

Example 4. The indicator function $\mathbb{I}_A$ is measurable if A is a measurable set.

Proposition 3. Given measurable maps

$$\Omega_1 \xrightarrow{f} \Omega_2 \xrightarrow{g} \Omega_3$$

the composition $g \circ f$ is also measurable.

Example 5. Given measurable $f, g : \Omega \to \mathbb{R}$, the maps $f \pm g$, $f \cdot g$, and $|f|$ are also measurable.

To define the Lebesgue integral, we'll start by defining it for simple functions (step functions with finite range), then extend it. We'll use it to integrate measurable functions $f : \Omega \to \mathbb{R}$ (where $\mathbb{R}$ is implicitly equipped with the Borel σ-algebra).

Take a simple function $h \in \mathbb{S}$ mapping $\Omega \to \mathbb{R}$ by

$$h = \sum_{i=1}^{N} c_i \mathbb{I}_{A_i}$$

for disjoint measurable A_i. Note that h is measurable since $x \mapsto c_i$ and $\mathbb{I}_{A_i}$ are measurable for all i. We can define the integral of h to be

$$I(h) := \sum_{i=1}^{N} c_i \; \mu(A_i),$$

which agrees with the intuition that the integral captures the area under the graph of a function (although this is now clearly generalizable to arbitrary dimensions, as long as Ω has a measure μ).

There's a problem here: suppose $h = 2\mathbb{I}_{A_1} - 3\mathbb{I}_{A_2}$ and $\mu(A_1) = \mu(A_2) = \infty$, then the integral evaluates to $I(h) = 2 \cdot \infty - 3 \cdot \infty$, which is undefined behavior so far. We have two main options to deal with this:

1. Only define the integral for simple maps with $|c_i| < \infty$ and $\mu(A_i) < \infty$ for all i.

2. Enforce $c_i \geq 0$ for all i (i.e. $h \geq 0$), then we can keep using ∞ without this problem.

We'll do the second option.

The representation of a simple function doesn't matter.

Definition 5. Denote the space of non-negative simple maps $\Omega \to \mathbb{R}$ by $\mathbb{S}^+$. For $h \in \mathbb{S}^+$, choose a representation

$$h = \sum_{i=1}^{N} c_i \mathbb{I}_{A_i}$$

with $c_i \geq 0$ for all i. Then the **Lebesgue integral** of h wrt μ is

$$\int_{\Omega} f \; d\mu := I(h) = \sum_{i=1}^{N} c_i \; \mu(A_i).$$

Now let $f : \Omega \to [0, \infty]$ be *any* measurable map. Then the Lebesgue integral of f is

$$\int_{\Omega} f \; d\mu = I(f) := \sup \left\{ I(h) \mid h \in \mathbb{S}^+, h \leq f \right\}.$$

A measurable map f is μ-**integrable** if $\int_{\Omega} f \; d\mu < \infty$.

Intuition here w/ sketch

A nice property of the Lesbesgue integral is that sets of measure 0 don't change its value.

Proposition 4. For $B \subset \Omega$ with $\mu(B) = 0$,

$$\int_{\Omega} f \; d\mu = \int_{\Omega - B} f \; d\mu$$

for all measurable $f \geq 0$.

Proof. Let $h \in \mathbb{S}^+$, then $h = \sum_{i=1}^{N} c_i \mathbb{I}_{A_i}$ for disjoint A_i. We can split this up as

$$h = \sum_{i=1}^{N} c_i \mathbb{I}_{(A_i \, \cap \, B^c)} + \sum_{i=1}^{N} c_i \mathbb{I}_{(A_i \, \cap \, B)}.$$

Since $0 \leq \mu(A_i \cap B) \leq \mu(B) = 0$ for all i, the integral of h is

$$\begin{aligned} I(h) &= \sum_{i=1}^{N} c_i \mu(A_i \cap B^c) + \sum_{i=1}^{N} c_i \mu(A_i \cap B) \\ &= \sum_{i=1}^{N} c_i \mu(A_i \cap B^c) + \sum_{i=1}^{N} 0 \\ &= I(h|_{\Omega - B}). \end{aligned}$$

□

This result is nice because it lets us care about properties that hold μ-a.e. instead of everywhere. In general, a property holds μ-a.e. if the set of points for which it doesn't hold has measure 0.

Proposition 5. Properties of the Lebesgue integral for general measurable maps:

1. Linearity: $I(\alpha f + \beta g) = \alpha I(f) + \beta I(g)$.
2. Monotonicity: if $f \leq g$ μ-a.e., then $I(f) \leq I(g)$.
3. If $f = g$ μ-a.e., then $I(f) = I(g)$.
4. $f = 0$ μ-a.e. $\iff$ $I(f) = 0$.

3 CONVERGENCE THEOREMS

The first two theorems in this section cover measurable maps $f \geq 0$. Lebesgue's dominated convergence theorem applies to any measurable map.

Theorem 1 (Monotone Convergence Theorem)**.** For measurable $\{f_n\}, f : \Omega \to [0, \infty]$ such that

1. $f_n \leq f_{n+1}$ for all n;
2. $\lim_{n\to\infty} f_n = f$ μ-a.e.;

then

$$\lim_{n\to\infty} \int_\Omega f_n \, d\mu = \int_\Omega f \, d\mu.$$

Proof. Since $f_n \leq f_{n+1}$ μ-a.e. for all n, and since $f_n \to f$ μ-a.e., we get that $f_n \leq f$ μ-a.e. Then by monotonicity of the integral,

$$\int_\Omega f_n \, d\mu \leq \int_\Omega f \, d\mu \text{ for all } n \implies \lim_{n\to\infty} \int_\Omega f_n \, d\mu \leq \int_\Omega f \, d\mu.$$

Proving the opposite inequality is a bit more involved. **Write this part down.** □

Define the limit inferior of a sequence of functions as follows:

$$\liminf_{n\to\infty} f_n(x) := \lim_{n\to\infty} \left(\inf_{k\geq n} f_k(x) \right).$$

Using this, we can derive Fatou's Lemma, which is a weak result but requires only weak conditions.

Theorem 2 (Fatou's Lemma)**.** Let $f_n : \Omega \to [0, \infty]$ be measurable. Then

$$\int_\Omega \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int_\Omega f_n \, d\mu.$$

Proof. Let $g_n := \inf_{k\geq n} f_k$, then $g_i \leq g_{i+1}$ for all i. Then by the MCT,

$$\int_\Omega \lim_{n\to\infty} g_n d\mu = \lim_{n\to\infty} \int_\Omega g_n d\mu = \liminf_{n\to\infty} \int_\Omega g_n d\mu \leq \liminf_{n\to\infty} \int_\Omega f_n d\mu.$$

Note that the second equality is true b/c whenever a limit exists, it's equal to the limit inferior/superior. The last inequality is true because $g_n \leq f_n$ for all n. □

Theorem 3 (Lebesgue's Dominated Convergence Theorem)**.** Let $f_n, f : \Omega \to \mathbb{R}$ be arbitrary measurable maps, with $f_n \to f$ pointwise μ-a.e. If there is some $g \in \mathcal{L}^1$ such that $|f_n| \leq g$ for all n (μ-a.e.), then $f_n, f \in \mathcal{L}^1$ for all n and

$$\lim_{n\to\infty} \int_\Omega f_n \, d\mu = \int_\Omega f \, d\mu.$$

4 CARATHÉODORY'S EXTENSION THEOREM

Definition 6. A family of subsets $\mathcal{A}$ of a set Ω is a **semiring of sets** if

1. $\varnothing \in \mathcal{A}$;
2. it is closed under pairwise intersections;
3. complements can be written as finite disjoint unions, i.e. for $A, B \in \mathcal{A}$, there are pairwise disjoint sets $S_1, \ldots, S_n \in \mathcal{A}$ such that $\bigcup_{i=1}^n S_i = A - B$.

Definition 7. Let $\mathcal{A}$ be a semiring of sets, then $\mu : \mathcal{A} \to [0, \infty]$ is a **premeasure** if

1. $\mu(\varnothing) = 0$;
2. $\mu\left(\bigcup_{i=1}^\infty A_i\right) = \sum_{i=1}^\infty \mu(A_i)$ for all pairwise disjoint $A_i \in \mathcal{A}$ *if* this union is in $\mathcal{A}$ (since $\mathcal{A}$ isn't necessarily a σ-algebra, this union might not actually be a member of $\mathcal{A}$).

Theorem 4 (Carathéodory's Extension Theorem)**.** Let $\mathcal{A} \subset P(\Omega)$ be a semiring of sets, and let μ be a premeasure on $\mathcal{A}$.

1. **Existence:** μ has an extension $\tilde{\mu} : \sigma(\mathcal{A}) \to [0, \infty]$ that's a measure. Here, being an extension means μ and $\tilde{\mu}$ agree on all $A \in \mathcal{A}$.
2. **Uniqueness:** if Ω can be covered by a sequence of sets in $\mathcal{A}$, each of which has finite measure under μ, then $\tilde{\mu}$ is unique (this condition says that μ must be σ-**finite**, so other equivalent characterizations of σ-finiteness would work here too).

An important application of this theorem is proving the existence of the Lebesgue measure. Let $\mathcal{A} = \{[a, b) \mid a, b \in \mathbb{R}, a \leq b\}$, which is a semiring of sets w/ $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R})$. Now define $\mu : \mathcal{A} \to [0, \infty]$ by $\mu([a, b)) = b - a$, which is a σ-finite premeasure on $\mathcal{A}$. Then by Carathéodory's Extension Theorem, there's a unique extension of μ that's a measure on $\mathcal{B}(\mathbb{R})$. This is the Lebesgue measure.

4.1 LEBESGUE-STIELTJES MEASURES

Let $F : \mathbb{R} \to \mathbb{R}$ be monotonically non-decreasing (can be discontinuous), and define

$$\mu([a, b)) := F(b^-) - F(a^-),$$

where $x^- := \lim_{\varepsilon \searrow 0} F(x - \varepsilon)$. Then by Carathéodory's Extension Theorem, there is a unique measure $\mu_F : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ such that $\mu_F([a, b)) = F(b^-) - F(a^-)$. This is the **Lebesgue-Stieltjes measure** for F. **Figure**

- If $F = \text{id}$, then μ_F is the Lebesgue measure.
- If F is a constant map, then $\mu_F = 0$.

- Fix α, and define F by
$$F(\alpha) = \begin{cases} 0 & x < \alpha, \\ 1 & x \geq \alpha. \end{cases}$$
Then
$$\mu_F([a,b)) = \begin{cases} 1 & \alpha \in [a,b), \\ 0 & \text{else.} \end{cases}$$
Note that for all $\varepsilon > 0$, $\mu_F([\alpha - \varepsilon, \alpha + \varepsilon)) = 1$. The Dirac measure at α also does this, so by uniqueness, $\mu_F = \delta_\alpha$.

- Let $F : \mathbb{R} \to \mathbb{R}$ be monotonically non-decreasing but also continuously differentiable, i.e. $F' : \mathbb{R} \to [0, \infty)$ is continuous (this means we don't have to worry about jumps or left limits anymore). Then
$$\mu_F([a,b)) = F(b) - F(a) = \int_a^b F'(x)\ dx.$$
More generally, we can define a **density map**
$$\begin{aligned} \mu_F : \mathcal{B}(\mathbb{R}) &\to [0, \infty] \\ A &\mapsto \int_A F'(x)\ dx. \end{aligned}$$

5 DECOMPOSITION THEOREMS

Let λ denote the Lebesgue measure on $\mathcal{B}(\mathbb{R})$. In this section, we'll be interested in other measures on $\mathcal{B}(\mathbb{R})$.

Definition 8. A measure μ is **absolutely continuous** wrt λ if

$$\lambda(A) = 0 \implies \mu(A) = 0;$$

i.e. μ is not "finer" than λ. Notation: $\mu \ll \lambda$.

Definition 9. A measure μ is **singular** wrt λ if there is some $N \in \mathcal{B}(\mathbb{R})$ such that

$$\lambda(N) = 0 \quad \text{and} \quad \mu(N^c) = 0.$$

Notation: $\mu \perp \lambda$.

Example 6. Let δ_α be the Dirac measure at α, then $\delta_\alpha \perp \lambda$. To see this, let $N = \{\alpha\}$, then $\lambda(N) = 0$ and $\delta^\alpha(N^c) = \delta_\alpha(\mathbb{R} - \{\alpha\}) = 0$.

Theorem 5. Let $\mu : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ be a σ-finite measure.

1. **Radon-Nikodym Theorem:** There exist 2 uniquely determined measures μ_{ac}, μ_s on $\mathcal{B}(\mathbb{R})$ such that
   - $\mu = \mu_{ac} + \mu_s$;
   - $\mu_{ac} \ll \lambda$ (absolutely continuous);
   - $\mu_s \perp \lambda$ (singular).
2. **Lebesgue's Decomposition Theorem:** There exists a measurable map (a **density map**) $h : \mathbb{R} \to [0, \infty)$ such that

$$\mu_{ac}(A) = \int_A h \, d\lambda$$

   for all $A \in \mathcal{B}(\mathbb{R})$. In other words, if a measure is absolutely continuous wrt λ and is σ-finite, then we can rewrite it as an integral wrt λ.

Lebesgue's decomposition theorem converts a very abstract concept (a measure) into something more concrete (a density, which is just a normal function). And by Radon-Nikodym, we can think about any general measure as having two orthogonal components: an "easy-to-use" component μ_{ac} that we can transform into a density, and a residual component μ_s that we know is singular.

Note 1. Nothing in this section was a special property of $\mathcal{B}(\mathbb{R})$. This all could've been written in terms of two arbitrary measure spaces $(\Omega_1, \mathcal{A}_1, \mu)$ and $(\Omega_2, \mathcal{A}_2, \lambda)$.

6 CHANGE OF VARIABLES

Given a measurable map $f : (\Omega_1, \mathcal{A}_1, \mu) \to (\Omega_2, \mathcal{A}_2)$, we can define the **pushforward measure** $f_*\mu$ on Ω_2 by

$$f_*\mu := \mu \circ f^{-1}.$$

Theorem 6 (Change of variables)**.** Suppose there are measurable maps

$$(\Omega_1, \mathcal{A}_1, \mu) \xrightarrow{f} (\Omega_2, \mathcal{A}_2) \xrightarrow{g} \mathbb{R}.$$

Then

$$\int_{\Omega_2} g \; d(f_*\mu) = \int_{\Omega_1} (g \circ f) \; d\mu.$$

Proof. **Put theorem about approximating measurable function w/ increasing sequence of simple functions.** Consider the case when $g = \mathbb{I}_A$ for some $A \in \mathcal{A}_2$. Then this holds by defintion:

$$\int_{\Omega_2} \mathbb{I}_A \; d(f_*\mu) = (f_*\mu)(A) = \mu(f^{-1}(A)) = \int_{\Omega_1} \mathbb{I}_{f^{-1}(A)} \; d\mu = \int_{\Omega_1} (\mathbb{I}_A \circ f) \; d\mu.$$

By linearity of the integral, we can extend this to the case where g is any simple function. Now let g be an arbitrary measurable map $\Omega_2 \to \mathbb{R}$, and take a sequence of simple functions $h_1 \leq h_2 \leq \cdots$ approximating it from below. Then $(h_1 \circ f) \leq (h_2 \circ f) \leq \cdots$ is an increasing sequence of functions approximating $h \circ f$ from below. Since the change of variables formula works for simple functions, we apply the MCT twice to get

$$\int_{\Omega_2} g \; d(f_*\mu) = \lim_{n\to\infty} \int_{\Omega_2} h_n \; d(f_*\mu) = \lim_{n\to\infty} \int_{\Omega_1} (h_n \circ f) \; d\mu = \int_{\Omega_1} (h \circ f) \; d\mu.$$

□

This is a more general form of the usual change of variables for 1-dimensional riemann integrals

$$\int_{\mathbb{R}} g(x) \; dx = \int_{\mathbb{R}} g(f(x) \cdot f'(x) \; dx.$$

To see this, consider the case when we have a surjective, strictly monotonically increasing, and continuously differentiable map $F : (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_F) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. By Carathêodory's extension theorem, we can determine the pushforward measure by where it sends intervals:

$$\begin{aligned}(F_*\mu_F)([a,b)) &= \mu\left(F^{-1}([a,b))\right) \\ &= \mu_F\left(\left[F^{-1}(a), F^{-1}(b)\right)\right) \qquad \text{(since } F \text{ is strictly monotonically increasing)} \\ &= \int_{F^{-1}(a)}^{F^{-1}(b)} F'(x) \; dx.\end{aligned}$$

Making a change of variables $y := F(x)$, this becomes

$$(F_*\mu)([a,b)) = \int_a^b dy = \lambda([a,b)).$$

Thus $F_*\mu_F$ is actually just the Lebesgue measure λ. Then in this case, the change of variables formula simplifies to

$$\int_{\mathbb{R}} g(x) \; dx = \int_{\mathbb{R}} g \; d(F_*\mu_F) = \int_{\mathbb{R}} (g \circ F) \; d\mu_F = \int_{\mathbb{R}} g(F(x)) \cdot F'(x) \; dx.$$