Measure Theory

How does one measure the “size” or “mass” of a set $S$ of elements from some space $X$? If $X = \mathbb{R}^n$ and $S$ is defined by a geometric shape, it makes sense to take the volume of $S$ as its measure. But in a more general sense, this won’t always work, especially when $X \neq \mathbb{R}^n$.

A measurable space is a pair $(X, M)$ of a set $X$ and a family $M$ of subsets of $X$. A measure space is a triplet $(X, M, \mu)$ where $\mu : M \rightarrow [0, \infty]$ is a function that measures the size of elements of $M$. Note that in general (unless $M$ is the power set on $X$), $\mu$ cannot measure every subset of $X$, it can only measure elements in the set $M$ of measurable subsets of $X$.

In order for $(X, M)$ to qualify as a measure space, $M$ must be a $\sigma$-algebra on $X$. That is, $M$ must satisfy the following axioms:

- $\emptyset \in M$ and $X \in M$.
- $M$ is closed under countable unions.
- $M$ is closed under complements.

By De’Morgan, it follows that $M$ must also be closed under intersections. Note that if $M$ has a countable basis, then $M$ is a topology, but not a very interesting one since every set is clopen.

Given a measurable space $(X, M)$, in order for $(X, M, \mu)$ to define a measure, $\mu$ must satisfy the measure axioms:

- $\mu(\emptyset) = 0$
- $\mu$ is countably additive. That is, for every countable collection of pairwise disjoint sets $\{E_i\}_{i=1}^\infty$ with $E_i \in M$, $\mu(\bigcup_{i=1}^\infty E_i) = \sum_{i=1}^\infty \mu(E_i)$.

The measure axioms imply several other properties, including monotonicity: If $E \subset F$, then $\mu(E) \leq \mu(F)$.

A set $N \in M$ is a null set if $\mu(N) = 0$. If $E \subset N$, then $E$ is a subnull set. A measure space is complete if it contains all its subnull sets. I.e., if $N$ is a subnull set, then $N \in M$.

If something is true everywhere except on a null set, then real analysis techniques cannot distinguish it from being true everywhere. To cover these corner cases, we say that some things are true almost everywhere. For example, $f = g$ almost everywhere if $f(x) = g(x)$ for all $x \in X \setminus N$ where $\mu(N) = 0$. 

If we relax the second measure axiom, allowing that \( \mu \) need only be subadditive (we must now also enforce that \( \mu \) is monotone as well), then we get an outer measure \( \mu^* \). Every measure corresponds to an outer measure on the entire space \( X \). Then the Caratheodory criterion states that a set \( E \) is \( \mu \)-measurable (that is, \( E \in M \), for \((X, M, \mu)\) where \( \mu = \mu^*|_M \)) if and only if
\[
\mu^*(A) = \mu^*(E \cap A) + \mu^*(E^C \cap A) \quad \text{for all } A \subseteq X.
\]

In the special case where \( X = \mathbb{R}^n \), \( M \) is the \( \sigma \)-algebra generated by boxes in \( \mathbb{R}^n \) (i.e., \( M \) is the coarsest \( \sigma \)-algebra containing every open or closed box in \( \mathbb{R}^n \), including single points), and \( \mu \) is defined as the unique measure that assigns the volume of a box as its measure, we get the standard measure space on \( \mathbb{R}^n \), denoted \((\mathbb{R}^n, \mathcal{L}, m)\) where \( m \) is the Lebesgue measure and \( \mathcal{L} \) are the Lebesgue measurable sets.

**Measurable Functions and Integration**

Let \( f : X \to \mathbb{R}^n \) be a function on a measure space \((X, M, \mu)\). Then \( f \) is measurable if \( f^{-1}(U) \in M \) for all open \( U \) in \( \mathbb{R}^n \) with the standard topology.

A special type of measurable function is the indicator function \( 1_E : X \to \{0, 1\} \) which returns \( 1_E(x) = 1 \) if \( x \in E \) and \( 1_E(x) = 0 \) otherwise. Another class of measurable functions are the simple functions, which are finite linear combinations of indicator functions:

\[
f_{\text{simp}}(x) = \sum_{i=1}^{n} c_i 1_{E_i}(x)
\]

where \( E_i \) are all measurable. We define the simple integral on simple functions by:

\[
\text{Simp} \int_X f_{\text{simp}}(x) d\mu = \sum_{i=1}^{n} c_i \mu(E_i).
\]

Note that the simple integral is both monotone (if \( f(x) \leq g(x) \) for all \( x \in X \) then \( \int_X f \, d\mu \leq \int_X g \, d\mu \)) and linear (\( \int_X (f + g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu \)).

In general, we define the unsigned integral for an unsigned measurable function \( f \) as:

\[
\int_X f \, d\mu = \sup_{g \leq f, \, g \in \text{Simp}(X)} \text{Simp} \int_X g \, d\mu.
\]

then we define the signed integral for a general measurable function \( f \):

\[
\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu
\]

where \( f^+ \) and \( f^- \) denote the positive and negative components of \( f \) respectively. We define integration even more generally for complex-valued functions by decomposing \( f \) into real and complex parts similarly as above.

If \( \int_X |f| d\mu < \infty \) then we say that \( f \) is absolutely integrable (or just integrable) and place it in the \( \mathcal{L}^1 \) class of functions. This induces a norm on \( f \), specifically, \( \|f\|_{L^1} = \int_X |f| \, d\mu \).
Integration Theory

Recall from elementary multivariable calculus, that we can exchange the order of integration for “most” functions, but there are some strange functions where one cannot. Let \( f : X \times Y \to \mathbb{C} \) where \((X, M_X, \mu_X)\) and \((Y, M_Y, \mu_Y)\) are measure spaces and \( X \times Y \) gets the product measure space (i.e., the coarsest measure space where all \( E \times F \) are measurable with \( E \in M_X \) and \( F \in M_Y \), and the measure is given by \( \mu_X(E)\mu_Y(F) \)).

Consider the integral: \( \int_{X \times Y} f \, dx \, dy \). We want to say that

\[
\int_{X \times Y} f \, dx \, dy = \int_X \int_Y f \, dy \, dx = \int_Y \int_X f \, dx \, dy
\]

but first we must show when it is true that

\[
\int_X \int_Y f \, dy \, dx = \int_Y \int_X f \, dx \, dy
\]

**Key question:** When can we exchange the order of a limit and an integral?

Answering this question tells us a lot, since integrals can be decomposed into limits of sums of simple functions.

**Monotone Convergence Theorem (MCT):**
Let \( \{f_n\}_{n=1}^{\infty} \) be measurable and unsigned non-decreasing functions (i.e., \( f_1 \leq f_2 \leq f_3 \leq \ldots \)). Then

\[
\lim_{n \to \infty} \int_X f_n \, d\mu = \int_X \lim_{n \to \infty} f_n \, d\mu.
\]

Note: The above has an analogue for nested sequences of sets

\[
E_n \uparrow \Rightarrow \lim_{n \to \infty} \mu(E_n) = \mu(\lim_{n \to \infty} E_n),
\]

and similarly for \( E_n \downarrow \).

**Dominated Convergence Theorem (DCT):**
Let \((X, M, \mu)\) be a measure space. Let \( \{f_n\}_{n=1}^{\infty} : X \to \mathbb{R} \) be measurable such that \( f_n \to f \) pointwise almost everywhere and suppose there exists an unsigned function \( g \in L^1 \) that dominates \( f_n \) almost everywhere (\(|f_n(x)| \leq g(x)\) except on a null set). Then

\[
\lim_{n \to \infty} \int_X f_n \, d\mu = \int_X \lim_{n \to \infty} f_n \, d\mu = \int f \, d\mu.
\]

**Tonelli’s Theorem and Fatou’s Lemma:**
Let \( \{f_i\}_{i=1}^{\infty} \) unsigned be measurable. Then

\[
\int_X \left( \sum_{n=1}^{\infty} f_n \right) \, d\mu = \sum_{n=1}^{\infty} \int_X f_n \, d\mu. \quad \text{(Tonelli)}
\]

and

\[
\int \liminf_{n \to \infty} f_n \, d\mu \leq \liminf_{n \to \infty} \int f_n \, d\mu. \quad \text{(Fatou)}
\]
Using the above Theorems, we arrive at the crux of Real Analysis:

First, we arrive at the conclusion that integration by parts is well-defined for all unsigned measurable functions. This is a direct consequence of Tonelli’s Theorem. Then, if \( f \) is signed, we get:

**Fubini’s Theorem:**
If \( f : M_X \times M_Y \to \mathbb{C} \) is absolutely integrable (in \( \mathcal{L}^1 \)) then the integral is well-defined and

\[
\int_{X \times Y} f \, dx \, dy = \int_X \int_Y f \, dy \, dx = \int_Y \int_X f \, dx \, dy.
\]

In conclusion, note that integration by parts works if \( f \) is *either* unsigned *or* absolutely integrable. In other words, integration by parts breaks down when \( f \) has both positive and negative parts and its integral is infinite.

**Littlewood’s Three Principles**

These require some interpretation, but give some intuition about the properties of measurable sets and functions. To interpret the word “nearly,” think “can be approximated to arbitrary precision by” or “except on a null set” as appropriate:

1. Every measurable set is nearly a finite union of intervals.
2. Every absolutely integrable function is nearly continuous.
3. Every pointwise convergent sequence of functions is nearly uniformly convergent.

It should be noted that these principles apply primarily to the special case of \( \mathbb{R}^n \) with the Lebesgue measure.

**Signed Measures**

On a final note, though we have only discussed unsigned measures: \( \mu : M \to [0, +\infty] \), we can also allow for signed measures \( \sigma : M \to \mathbb{R} \). However, every signed measure \( \sigma \) can be decomposed into the difference between two unsigned measures \( \sigma^+ \) and \( \sigma^- \). That is,

\[
\sigma = \sigma^+ - \sigma^-.
\]