Lebesgue Integration

Objective: To understand the basics of measure and integration, and apply them to inner products on L_2 spaces.

Introduction: Measure:

A measure is, roughly speaking, a way of describing the size of a set. More precisely, if you have a set S, it is often useful to be able to compare the size of subsets of S, in some sense.

There is already a notion of size for sets, the **cardinality** (number of elements), and this is indeed a measure, but it's not always a useful one, particularly when dealing with infinite sets. For instance, if $S = \mathbb{R}$, we can consider the two subsets $[0, 1]$ and $[0, 2]$; these both have the same cardinality (the "cardinal of the continuum"), so in terms of cardinality they are the same size; but thinking of them as line segments, we are perfectly comfortable saying that $[0, 2]$ is twice as long as $[0, 1]$.

So a measure on a set S is a notion of "size" for the subsets of S, which can be more refined than simply the number of elements each subset contains. Examples of situations where we want a different notion of the size of a set than its cardinality come in geometry, where length (for subsets of $\mathbb{R}$), area (for subsets of $\mathbb{R}^2$), and volume (for subsets of $\mathbb{R}^3$).

Another example is in probability theory. Given an experiment (or some occurence with a defined set of outcomes), the **sample space** is defined to be the set of all possible outcomes, an **event** is a subset of the sample space, and probability is a measure which assigns to each event how likely it is.

For instance, when rolling a normal, 6-sided die, the sample space is the set of possible rolls, $S = \{1, 2, 3, 4, 5, 6\}$, and examples of events would be the event of rolling a 1, which is $\{1\}$, or the event of rolling an even number, which is $\{2, 4, 6\}$. The advantage of working with subsets of the sample space instead of just individual outcomes is that it lets us group together several outcomes as members of the same event. For a finite sample space like this one, we can just use the number of outcomes in an event as its size, so because $\{2, 4, 6\}$ has three times as many elements as $\{1\}$, we say rolling an even number is three times as likely as rolling a 1. Rather than the actual sizes of the events, we tend to use the size of the event divided by the total number of possible outcomes, so that the total probability of the whole sample space is 1. So $\{2,4,6\}$ has probability $\frac{3}{6}=\frac{1}{2}$, while $\{1\}$ has probability $\frac{1}{6}$.

For a continuous random variable, such as the shortest distance from one tree to another in a forest, we cannot simply use the size of a set (divided by the size of the sample space) as the probability. Since the distance from one tree to the next can be any positive real number, the sample space is infinite; so we need a measure of how likely a certain subset of the sample space (a certain range of possible distances) is to occur.

We could use the idea of length, and say, for instance, that the probability of the nearest tree being between a and b cm away is $b-a$ (*i.e.*, the probability of the event $[a,b]$ in the sample space $\mathbb{R}_+$ is the length $b-a$), but this is unlikely to accurately model real forests. Trees are not usually right next to each other, so the probability assigned to the interval $[0,100]$, say, should be quite small. So we need some alternative way of assigning "size" (probability) to subsets of the sample space.

So there are several different contexts in which we might need to describe the "size" of a subset of some set, and several different notions of size we might want to consider. It is therefore worth considering what general properties any sensible notion of "size" ought to have. This is the goal of Measure Theory.

The basic properties settled on for a measure is that it should be give a non-negative real number or infinity for each set (sets can have size 0, or be infinitely big, or have any size in between, but cannot be negative), it should give 0 for the empty set (the empty set should have size 0), and it should be "countably additive over pairwise disjoint sets," meaning that if $A_1, A_2, \ldots$ are sets such that A_i and A_j have no elements in common when $i \neq j$, then the measure of the union should be the sum of the individual measures (if you put together a bunch of separate sets, their combined size should be the sum of their individual sizes).

The problem is that it can be shown that there is, in general, no consistent way to do this for all subsets of a set! We will show this in the exercises. It is possible to define really weird subsets of fairly nice sets like the real numbers, such that there is no sensible way to define the "size" of such a subset. However, such subsets are unlikely to arise in practical problems; so the solution is simply to restrict the measure to "nice" subsets.

Theory: σ-Algebras and Measure Theory:

Given a set S, a σ-**algebra** on S is a collection Σ of subsets of S (called Σ-sets), satisfying the following three properties:

1. $S \in \Sigma$,
2. if $A \in \Sigma$, then $S \setminus A \in \Sigma$ (the complement of a Σ-set is a Σ-set)
3. If $A_1, A_2, \ldots \in \Sigma$, then
$$\bigcup_{n=1}^{\infty} A_n \in \Sigma$$
(the union of countably many Σ-sets is a Σ-set).

The idea is that Σ-sets are "nice" subsets of S, on which we will be able to define our measure. Subsets of S that aren't in Σ are "bad" subsets, and we won't be able to measure their size. The axioms of a σ-algebra then are saying that S is a "nice" set (and we will be able to say how big S is), the complement of a nice set is nice (and we will define the size of $S \setminus A$ to be the size of S minus the size of A), and a union of (countably many) nice sets is nice (so we can't put nice sets together and make a bad set, unless we use a truly enormous number of nice sets).

Note that $\varnothing \in \Sigma$, by 1 and 2, and the intersection of countably many Σ-sets is a Σ-set, by 2 and 3 (with de Morgan's Laws).

If S is a set and Σ is a σ-algebra on S, a **measure** on the pair (S, Σ) is a function $\mu : \Sigma \to [0, \infty]$ (*i.e.*, μ takes a Σ-set as input and gives a non-negative real number or "infinity" as output), such that:

1. $\mu(\varnothing) = 0$ (the measure of the empty set is 0),
2. if $A_1, A_2, \ldots \in \Sigma$ and $A_i \cap A_j = \varnothing$ for all $i \neq j$, then
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu\left(A_n\right),$$
where any sum with ∞ as a summand is taken to have the value ∞.

Lebesgue Measure:

There are lots of different measures, on lots of different sets. The one of interest to us will be the **Lebesgue measure** on $\mathbb{R}$ (there is also a Lebesgue measure on $\mathbb{R}^n$ for each n, but we will only consider $n = 1$). To define the Lebesgue measure, we first need to define a σ-algebra to be the domain of the measure function. In fact, we start by defining an "outer measure" (like a measure, but not quite), use that to define the Lebesgue σ-algebra, then refine the outer measure to define the actual Lebesgue measure.

So define the **Lebesgue outer measure** λ^+ as follows. For any subset E of $\mathbb{R}$, we approximate E from above by a countable sequence of open intervals; that is, we find intervals (a_n, b_n) such that E is contained in the union of these intervals:

$$E \subseteq \bigcup_{n=1}^{\infty} (a_n, b_n).$$

Then the length of each interval is given by $b_n - a_n$, so we can add up all these lengths to get an *over*estimate of the size of E:

$$\lambda^+(E) \leq \sum_{n=1}^{\infty} (b_n - a_n) \, .$$

Note that we haven't assumed that the intervals (a_n, b_n) are pairwise disjoint; just that between them they cover E. Now we consider doing this over all possible countable collections of open intervals that would cover E, getting a wide range of overestimate of $\lambda^+(E)$; we then take E to be the greatest lower bound (infimum) of these overestimates:

$$\lambda^+(E) = \inf \left\{ \sum_{n=1}^{\infty} (b_n - a_n) : E \subseteq \bigcup_{n=1}^{\infty} (a_n, b_n) \right\} .$$

We then say that a set E is **Lebesgue-measurable** if for any other set $A \subseteq \mathbb{R}$,

$$\lambda^+(A) = \lambda^+ (A \cap E) + \lambda^+ (A \cap (\mathbb{R} \setminus E)) \, .$$

So to say that E is Lebesgue-measurable is to say that when finding the outer measure of any set A, you can split A up into the part inside E and the part outside E, find their outer measures separately, then add them together. The **Lebesgue σ-algebra** is the set of all Lebesgue-measurable subsets of $\mathbb{R}$. For a Lebesgue-measurable set E, the Lebesgue measure λ is defined to simply be the same as the outer measure:

$$\lambda(E) = \lambda^+(E).$$

Theory: Lebesgue Integration:

Let $f : [a, b] \to \mathbb{R}$ be a non-negative function (so $f(x) \geq 0$ for all x). To integrate f by Riemann's method, we would subdivide the interval $[a, b]$ on the x-axis into subintervals, draw a rectangle over each subinterval, sum the areas of the rectangles, then let n tend to infinity. An alternative approach is to subdivide the y-axis, for instance into the intervals $[0, 2^{-n}]$, $[2^{-n}, 2 \times 2^{-n}]$, ..., $[(k-1)2^{-n}, k2^{-n}]$, ..., and then define $m_{n,k}(f)$ to be $\lambda\left(f^{-1}[k2^{-n}, (k+1)2^{-n}]\right)$, the Lebesgue measure of the set of points x such that $k2^{-n} \leq f(x) \leq (k+1)2^{n}$. We can then approximate the area under the graph of f by

$$\sum_{k=1}^{\infty} k2^{-n} m_{n,k}(f).$$

The limit of this expression as n tends to infinity is then called the **Lebesgue integral** of f. Note that the value of the integral can be ∞, if the limit as $n \to \infty$ does not exist; however, as n increases the value of the above sum increases (see exercises), so the only way the limit can fail to exist is if it grows without bound. This justifies saying the integral of f is infinite if the limit does not exist.

There is a potential problem with this. Since not all subsets of $\mathbb{R}$ are Lebesgue-measurable (and the preimage of an interval under a function can be a very complicated set), we can't guarantee that the numbers $m_{n,k}(f)$ are well-defined. As such, we must restrict to **measurable functions**—functions f such that the preimage of an interval is always a Lebesgue-measurable set. We will show in the exercises that most functions we are interested in are indeed measurable.

We can then extend this definition to cover non-positive measurable functions straightforwardly; if $f(x) \leq 0$ for all x, then $-f(x) \geq 0$ and $-f(x)$ is measurable (exercise), so we integrate $-f$ by the above approach and define

$$\int_a^b f(x)\,\mathrm{d}x = -\int_a^b -f(x)\,\mathrm{d}x.$$

For a measurable function which takes both positive and negative values, we split it into the positive part $f_+(x)$ and the negative part $f_-(x)$, and define

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b f_+(x)\,\mathrm{d}x + \int_a^b f_-(x)\,\mathrm{d}x,$$

unless this gives us the expression "$\infty - \infty$," in which case we cannot assign a sensible value to the integral.

So we define a function f to be **Lebesgue-integrable** on $[a, b]$ if f is measurable and we do not have the "$\infty - \infty$" problem when integrating.

Application: Inner Products on L_2-spaces:

Let I be an interval in the real line (of finite or infinite length—in particular, $I = \mathbb{R}$) is allowed. Define $L_2(I)$ to be the set of complex-valued measurable functions f on I such that

$$\int_I f(x)\overline{f(x)}\,\mathrm{d}x < \infty.$$

We often call such a function "square-integrable."

There is a problem here: we have only defined what it means for a function to be measurable if it is real-valued. For a function f from a subset of $\mathbb{R}$ to $\mathbb{C}$, we say that f is measurable if the preimage of any rectangle is measurable. That is, if for any a, b, c, and d,

$$\{x : a \le \mathrm{Re}(f(x)) \le b \text{ and } c \le \mathrm{Im}(f(x)) \le d\}$$

is measurable.

We will show that $L_2(I)$ is a vector space over $\mathbb{C}$, and has a complex inner product given by

$$\langle f \mid g\rangle = \int_I f(x)\overline{g(x)}\,\mathrm{d}x.$$

There are several things here we have to prove, and we shall do so in the exercises. Firstly we must show that $L_2(I)$ is a vector space over $\mathbb{C}$. This means showing that there are well-defined operations of addition and scalar multiplication (by complex scalars), and that these operations satisfy certain axioms. It is obvious what the definitions of $f + g$ and αf should be for $f, g \in L_2(I)$ and $\alpha \in \mathbb{C}$, but less obvious that $f + g$ and αf will still be in $L_2(I)$; if two square-integrable functions can be added to give a non-square-integrable function, or if one can be scaled to give a non-square-integrable function, then the obvious addition and scalar multiplication operations will not be well-defined on $L_2(I)$.

So we must first check that if f and g are square-integrable and $\alpha \in \mathbb{C}$, then $f + g$ and αf are square-integrable. Then we must check that these operations satisfy the axioms of a vector space; but this step is easy—just go through the axioms one-by-one and check them. Then we have that $L_2(I)$ is a $\mathbb{C}$-vector space, and need to check the inner product. Again, this involves two steps: first checking that the integral converges, and secondly checking that the axioms of a complex inner product are satisfied; again, the first of these is the harder step, and the second is just a routine verification.

Exercises:

1. In this exercise, we prove that the Lebesgue-measurable sets form a σ-algebra and that the Lebesgue measure is indeed a measure.

   (a) First we establish some properties of the Lebesgue outer measure. Let λ^+ be the Lebesgue outer measure:

   $$\lambda^+(E) = \inf\left\{\sum_{n=1}^{\infty}(b_n - a_n) : E \subseteq \bigcup_{n=1}^{\infty}(a_n, b_n)\right\}.$$

   Show that:

   i. $\lambda^+(\varnothing) = 0$;

   ii. If $A \subseteq B \subseteq \mathbb{R}$, then $\lambda^+(A) \leq \lambda^+(B)$;

   iii. If $A_1, A_2, \ldots$ are subsets of $\mathbb{R}$, then

   $$\lambda^+\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \lambda^+ A_n.$$

   (b) Next we show that the Lebesgue-measurable sets form a σ-algebra. Recall that a set E is Lebesgue-measurable if for any other set $A \subseteq \mathbb{R}$,

   $$\lambda^+(A) = \lambda^+\left(A \cap E\right) + \lambda^+\left(A \cap (\mathbb{R} \setminus E)\right).$$

   Show that:

   i. $\mathbb{R}$ is Lebesgue-measurable.

   ii. If E is Lebesgue-measurable, then so is $\mathbb{R} \setminus E$.

   iii. If $E_1, E_2, \ldots$ are Lebesgue-measurable, then so is

   $$\bigcup_{n=1}^{\infty} E_n.$$

   (c) Finally we show that, when restricted to the σ-algebra of Lebesgue-measurable sets, the Lebesgue outer measure is actually a measure. We already know that $\lambda^+(\varnothing) = \varnothing$, so we only need to check that if $E_1, E_2, \ldots$ are Lebesgue-measurable *and pairwise disjoint*, then

   $$\lambda^+\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \lambda^+(E_n).$$

2. In this exercise we show that Lebesgue integration for non-negative, measurable functions is well-defined. Recall that for such a function the Lebesgue integral over an interval I is defined to be

$$\int_I f(x)\,\mathrm{d}x = \lim_{n\to\infty} \sum_{k=1}^{\infty} k 2^{-n} m_{n,k}(f),$$

where $m_{n,k}(f) = \lambda\left(f^{-1}[k2^{-n},(k+1)2^{-n}]\right)$ is the Lebesgue measure of the preimage of the interval $[k2^{-n},(k+1)2^{-n}]$ under f.

(a) Show that

$$\begin{aligned} f^{-1}[k2^{-n},(k+1)2^{-n}] = f^{-1}&[(2k)2^{-(n+1)},(2k+1)2^{-(n+1)}] \\ &\cup f^{-1}[(2k+1)2^{-(n+1)},(2k+2)2^{-(n+1)}]. \end{aligned}$$

(b) Hence show that

$$k2^{-n}m_{n,k}(f) \le (2k)2^{-(n+1)}m_{n+1,2k} + (2k+1)2^{-(n+1)}m_{n+1,2k+1}.$$

(c) Hence show that

$$\sum_{k=1}^{\infty} k2^{-n}m_{n,k}(f) \le \sum_{k=1}^{\infty} k2^{-(n+1)}m_{n+1,k}(f).$$

(d) Hence conclude that as $n \to \infty$,

$$\sum_{k=1}^{\infty} k2^{-n}m_{n,k}(f)$$

converges either to a finite limit or to infinity, and hence the Lebesgue integral of a non-negative, measurable function is well-defined.

3. In this question, we show that various functions $\mathbb{R} \to \mathbb{R}$ are measurable. Show that:

(a) If f is continuous then f is measurable;

(b) If f and g are measurable and α is a constant, then so are $f(x)+g(x)$, $f(x)g(x)$, $f(g(x))$, and $\alpha f(x)$.

(c) If S is a Lebesgue-measurable subset of $\mathbb{R}$, then the indicator function $i_S(x)$ (1 when $x \in S$, 0 otherwise) is measurable.

(d) If f is measurable, then f_+ and f_- are measurable, where

$$f_+(x) = \frac{f(x)+|f(x)|}{2}, \quad f_-(x) = \frac{f(x)-|f(x)|}{2}.$$

Similar results can be shown for complex-valued functions. In particular, if $f : I \to \mathbb{C}$ is a complex-valued, measurable function on an interval I, then $f(x)\overline{f(x}$ is measurable, since it is the product of f with $g(f)$, where g is the continuous function complex-conjugation. This means the integral in the definition of an L_2 function is well-defined.

4. In this question we prove that $L_2(I)$ is an inner product space (for I an interval in $\mathbb{R}$).

   (a) First we show that $L_2(I)$ is a $\mathbb{C}$-vector space. The results of the previous question, on measurable functions, will be useful here.

      i. Show that if $f \in L_2(I)$ (so f is measurable and $f(x)\overline{f(x)}$ has finite integral), then $\alpha f \in L_2(I)$ for any constant $\alpha \in \mathbb{C}$.

      ii. Show that for any complex numbers z and w, $|z+w|^2 \leq 2|z|^2 + 2|w|^2$.

      iii. Hence show that, for f and g in $L_2(I)$, $f + g \in L_2(I)$.

      iv. Check the axioms of a vector space to complete the proof that $L_2(I)$ is a $\mathbb{C}$-vector space.

   (b) Next we show that the inner product on $L_2(I)$ is well-defined. It follows from our results on measurable functions (assuming the definition of measurability for complex-valued functions) that $f(x)\overline{g(x)}$ is measurable, so we just need to show that it has a finite integral, then check the axioms of an inner product.

      i. Show that for any complex numbers z and w

      $$2|z||w| \leq |z|^2 + |w|^2.$$

      Hint: consider $(|z| - |w|)^2$.

      ii. Hence show that for $f, g \in L_2(I)$,

      $$\int_I f(x)\overline{g(x)}\,\mathrm{d}x < \infty$$

      and therefore the inner product is well-defined.

      iii. Check the axioms of a complex inner product (linearity in the first argument, conjugate-symmetry, and positive-definiteness) to complete the proof that $L_2(I)$ is an inner product space.

5. In this question we show that not all subsets of $\mathbb{R}$ are Lebesgue measurable. Therefore it really is necessary to introduce the machinery of σ-algebras to work with measures.

(a) First we must prove a property of the Lebesgue measure called **translation-invariance**. This says that if $E \subseteq \mathbb{R}$ is a measurable set and $x \in \mathbb{R}$, then $x + E = \{x + e : e \in E\}$ is measurable and $\lambda(E) = \lambda(x + E)$—so adding a fixed amount onto every element of a set doesn't change its size.

i. First show that if A is any subset of $\mathbb{R}$ and (a_1, b_1), (a_2, b_2), $\dots$ is a (countable) sequence of intervals such that

$$A \subseteq \bigcup_{n=1}^{\infty} (a_n, b_n),$$

then

$$x + A \subseteq \bigcup_{n=1}^{\infty} (a_n + x, b_n + x).$$

So we can turn a cover of A by open intervals into a cover of $x + A$, simply by translating the open intervals by x.

ii. Hence show that for any A we have $\lambda^+(x + A) = \lambda^+(A)$.

iii. Hence show that if E is Lebesgue measurable, so is $x + E$. Hint: show that $A \cap (x + E) = x + ([(-x) + A] \cap E)$ and $\mathbb{R} \setminus (x + E) = x + (\mathbb{R} \setminus e)$.

iv. Hence conclude that λ is translation-invariant.

(b) Now we construct a set, called a **Vitali set**, which we will show is not measurable. We start with the rationals $\mathbb{Q}$, and consider translations $x + \mathbb{Q}$.

i. Show that for any two real numbers x and y, if $x - y \in \mathbb{Q}$, then $x + \mathbb{Q} = y + \mathbb{Q}$, and if $x - y \in \mathbb{R} \setminus \mathbb{Q}$, then $(x + \mathbb{Q}) \cap (y + \mathbb{Q}) = \varnothing$.

ii. Show that for any real x, $(x + \mathbb{Q}) \cap [0, 1] \neq \varnothing$ (*i.e.*, there is always an element of $x + \mathbb{Q}$ between 0 and 1).

iii. Choose exactly one element from each $(x + \mathbb{Q}) \cap [0, 1]$ for $x \in \mathbb{R}$, and let V be the set containing your choices (and nothing else). So V is a subset of $[0, 1]$ and for any two elements x and y of V, $x + \mathbb{Q} \neq y + \mathbb{Q}$. Note that the existence of V depends on a mildly controversial axiom of set theory called the Axiom of Choice. Show that for any $q \in \mathbb{Q}$, $q + V$ contains exactly one rational number.

(c) Let $q_1, q_2, \dots$ be an enumeration of the rational numbers in $[-1, 1]$ (so every rational number in $[-1, 1]$ appears exactly once in this list; this is possible because the rationals are countable). Let $V_i = V + q_i$ for each i. Show that $V_i \cap V_j = \varnothing$ for $i \neq j$.

(d) Now we show that

$$[0, 1] \subseteq \bigcup_{i=1}^{\infty} V_i \subseteq [-1, 2].$$

i. For $r \in [0,1]$, there is a unique element v of $r + \mathbb{Q}$ in V (by definition of V). Show that $v - r \in \mathbb{Q} \cap [-1,1]$.

ii. Hence there is some i such that $q_i = v - r$. Conclude that $r \in V_i$.

iii. Hence conclude that

$$[0,1] \subseteq \bigcup_{i=1}^{\infty} V_i.$$

iv. Use the fact that $V \subseteq [0,1]$ (by definition) to show that

$$\bigcup_{i=1}^{\infty} V_i \subseteq [-1,2].$$

(e) Now assume for a contradiction that V is Lebesgue measurable. By translation-invariance, so is each V_i, and $\lambda(V_i) = \lambda(V)$. Explain why it follows that

$$\lambda([0,1]) \leq \sum_{i=1}^{\infty} \lambda(V) \leq \lambda([-1,2]).$$

(f) Deduce a contradiction, and hence conclude that V is not in fact Lebesgue measurable.