MAU22200 Lecture 35

John Stalker

Trinity College Dublin

25 January 2022

What is measure theory?

This semester of MAU22200 covers measure and integration. Measure theory attempts to unify a number of closely related concepts.

Examples include cardinality of sets, probabilities of events, areas of subsets of the plane, volumes of subsets of three dimensional space, etc. In this module we're primarily interested in the last two examples, and their generalisation to n dimensions, but it's helpful to keep the others in mind.

Cardinality The cardinality of a set is non-negative, i.e. $\#A \geq 0$. Cardinality is monotone, i.e. if $A \subseteq B$ then $\#A \leq \#B$. The cardinality of the empty set is 0, i.e. $\#\varnothing = 0$. Cardinality is additive, i.e. $\#(A \cup B) + \#(A \cap B) = \#A + \#B$. In particular, if $A \cap B = \varnothing$ then $\#(A \cup B) = \#A + \#B$.

Probability Suppose Ω is the set of possible outcomes of an experiment and $\mathbf{P}(A)$ is the probability that the outcome belongs to some subset $A \subseteq \Omega$. Probability is non-negative, i.e. $\mathbf{P}(A) \geq 0$.

What is measure theory? (Continued)

Probability is monotone, i.e. if $A \subseteq B$ then $\mathbf{P}(A) \leq \mathbf{P}(B)$. The probability of an outcome in the empty set is 0, i.e. $\mathbf{P}(\varnothing) = 0$. Probability is additive, i.e. $\mathbf{P}(A \cup B) + \mathbf{P}(A \cap B) = \mathbf{P}(A) + \mathbf{P}(B)$. In particular, if $A \cap B = \varnothing$ then $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$.

Area The area of a subset of the plane is non-negative, i.e. $\alpha(A) \geq 0$. Area is monotone, i.e. if $A \subseteq B$ then $\alpha(A) \leq \alpha(B)$. The area of the empty set is 0, i.e. $\alpha(\varnothing) = 0$. Area is additive, i.e. $\alpha(A \cup B) + \alpha(A \cap B) = \alpha(A) + \alpha(B)$. In particular, if $A \cap B = \varnothing$ then $\alpha(A \cup B) = \alpha(A) + \alpha(B)$.

Volume, of course is similar.

There are differences, of course. Probabilities are finite, but cardinalities, areas and volumes can be infinite, for example. The empty set is the *only* set with cardinality 0, but it's not the only set with area 0 or volume 0. But there's enough similarity to try to build a general theory.

Why is measure theory hard?

Let's try to build an axiomatic theory of volume. It seems reasonable to suppose that

- Every set in three dimensional space has a non-negative, but possibly infinite, volume.
- Volume is monotone.
- The volume of the empty set is 0.
- Volume is (finitely) additive.
- Congruent sets have the same volume.
- Balls of radius r have volume $\frac{4}{3}\pi r^3$.

We might expect to need more axioms, but this seems like a good start. In fact, we've already assumed too much. These axioms are logically inconsistent!

Banach-Tarski

Theorem (Banach-Tarski): There are sets E_1, E_2, E_3, E_4, E_5 and F_1, F_2, F_3, F_4, F_5 such that

- E_i is congruent to F_i for each i,
- $E_i \cap E_j = \varnothing$ and $F_i \cap F_j = \varnothing$ when $i \neq j$,
- $E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5$ is a ball of radius 1.
- $F_1 \cup F_2 \cup F_3 \cup F_4 \cup F_5$ is the union of two balls of radius 1, which do not intersect.

If the axioms above are satisfied then E_i and F_i have the same volume but also the sum of the volumes of the F_i is twice the the sum of the volumes of the E_i. Note that these sums are finite and positive, so we have a contradiction.

How to escape the contradiction?

There are two lessons to be drawn from Banach-Tarski:

- We need to drop at least one of our axioms.
- No everything which is obvious is true.

The axiom we drop is the first one, that every subset of three dimensional space has a non-negative, but possibly infinite, volume. We don't drop non-negativity, we just don't try to assign a volume to *every* subset. Some sets are just too ugly to assign a volume to.

There are other options. One would be to drop the Axiom of Choice. That's not a very popular option.

The only way to deal with the second problem is to be very careful. Don't trust your intuition!

Integration

In fact this module isn't primarily about measure theory at all though. It's mostly about integration, but you need measure theory for a proper discussion of integration.

You may have seen integrals introduced informally as the area under the graph. That's not actually how we'll end up defining them, but it will be a consequence of the definition. Area is an example of a measure, specifically of Lebesgue measure.

One advantage of doing things in a more general context is that you can avoid repeating arguments. For example, the same theorems (Fubini-Tonelli) will govern exchanging the order of two two sums, exchanging a sum and an integral, or exchanging two integrals. There's also a corresponding statement for probability theory, though we don't have the language to express it.

Fubini-Tonelli Counterexample 1

Theorems have hypotheses. We can't always exchange the order of sums or integrals. Let

$$\alpha_{i,j} = \begin{cases} 1 & \text{if } i = j, \\ -1 & \text{if } i = j+1, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\sum_{j=0}^{\infty} \alpha_{i,j} = \begin{cases} 1 & \text{if } i = 0, \\ 0 & \text{if } i \neq 0. \end{cases} \qquad \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \alpha_{i,j} = 1.$$

$$\sum_{i=0}^{\infty} \alpha_{i,j} = \begin{cases} -1 & \text{if } j = -1, \\ 0 & \text{if } j \neq -1. \end{cases} \qquad \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \alpha_{i,j} = 0.$$

So $\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \alpha_{i,j} \neq \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \alpha_{i,j}$.

Fubini-Tonelli Counterexample 2

The same can happen for integrals.

$$\begin{aligned}
\int_0^1 \int_0^1 \frac{xy(x^2-y^2)}{(x^2+y^2)^3}\,dy\,dx &= \int_0^1 \int_0^1 \frac{\partial}{\partial y}\left(\frac{1}{2}\frac{xy^2}{(x^2+y^2)^2}\right)\,dy\,dx \\
&= \int_0^1 \left(\frac{1}{2}\frac{x}{(1+x^2)^2}\right)\,dx \\
&= \int_0^1 \frac{\partial}{\partial x}\left(-\frac{1}{8}\frac{1-x^2}{1+x^2}\right)\,dy = -\frac{1}{8}.
\end{aligned}$$

$$\begin{aligned}
\int_0^1 \int_0^1 \frac{xy(x^2-y^2)}{(x^2+y^2)^3}\,dx\,dy &= \int_0^1 \int_0^1 \frac{\partial}{\partial x}\left(-\frac{1}{2}\frac{x^2y}{(x^2+y^2)^2}\right)\,dx\,dy \\
&= \int_0^1 \left(-\frac{1}{2}\frac{y}{(1+y^2)^2}\right)\,dy \\
&= \int_0^1 \frac{\partial}{\partial y}\left(\frac{1}{8}\frac{1-y^2}{1+y^2}\right)\,dy = \frac{1}{8}.
\end{aligned}$$

Riemann-Jordan vs Lebesgue Integration

Ideally you would like a theory of integration in $\mathbf{R}^n$

- in which the hypotheses are weak enough to cover all the important examples,
- but not so weak as to include the counterexamples, so the theorems are actually true,
- and the definitions and proofs are fairly straightforward.

There is no known theory which satisfies all of these requirements. Riemann integration, which you learned in first year, and its higher dimensional analogue due to Jordan, satisfy the second and (sort of) the third, but not the first. Lebesgue integration, which you'll learn this semester, satisfies the second and (sort of) the first, but not the third.