MAU 22200 Week 9 Lecture 1

John Stalker

Trinity College Dublin

29 March 2021

Reading for this week

For this week, read the introduction to Section 1.5 (Modes of Convergence) and skim

- Subsection 1.5.1 (Uniqueness)
- Subsection 1.5.2 (The case of a step function)
- Subsection 1.5.3 (Finite measure spaces)
- Subsection 1.5.4 (Fast convergence)
- Subsection 1.5.5 (Domination and uniform integrability)

There are some results there which are useful in other modules, but the rest of the book only requires two exercises, 1.5.10 and 1.5.13, and it only requires them once, in the proof of Theorem 1.6.40 (the Second Fundamental Theorem of Calculus for absolutely continuous functions).
Also, read the introduction to Section 1.6 (Differentiation Theorems).

Miscellaneous comments

I'm not going to assign new exercises this week. You can use this week to get caught up if you're behind.

The published version of Terry's book and the version available online are nearly identical, but not quite. I'd stopped checking that the numbering of the exercises matched, but in Section 1.4 there were some differences. Sorry for any resulting confusion.

The numbering for Sections 1.5, 1.6. and 1.7 appears to match.

Modes of convergence

Section 1.5 discusses modes of convergence of sequences of functions. There are a lot of them, some old and some new.

- $f_n \to g$ pointwise if for all $x \in X$ and all $\epsilon > 0$ there is an $N > 0$ such that for all $n > N$ one has $|f_n(x) - g(x)| < \epsilon$.
- $f_n \to g$ uniformly if for all $\epsilon > 0$ there is an $N > 0$ such that for all $x \in X$ and all $n > N$ one has $|f_n(x) - g(x)| < \epsilon$.
- $f_n \to g$ pointwise almost everywhere if there is an $E \in \mathcal{B}$ with $\mu(E) = 0$ such that for all $x \in X \backslash E$ and all $\epsilon > 0$ there is an $N > 0$ such that for all $n > N$ one has $|f_n(x) - g(x)| < \epsilon$.
- $f_n \to g$ uniformly almost everywhere, a.k.a. in the L^∞ norm, if there is an $E \in \mathcal{B}$ with $\mu(E) = 0$ such that for all $\epsilon > 0$ there is an $N > 0$ such that for all $x \in X \backslash E$ and all $n > N$ one has $|f_n(x) - g(x)| < \epsilon$.

Modes of convergence, continued

- $f_n \to g$ almost uniformly if for all $\delta > 0$ there is an $E \in \mathcal{B}$ with $\mu(E) < \delta$ such that for all $\epsilon > 0$ there is an $N > 0$ such that for all $x \in X \backslash E$ and all $n > N$ one has $|f_n(x) - g(x)| < \epsilon$.
- $f_n \to g$ in the L^1 norm if for all $\epsilon > 0$ there is an $N > 0$ such that for all $n > N$ one has $\int_X |f_n(x) - g(x)| \, d\mu < \epsilon$.
- $f_n \to g$ in measure if for all $\delta > 0$ and all $\epsilon > 0$ there is an $N > 0$ such that for all $n > N$ one has $\mu(\{x \in X : |f_n(x) - g(x)| > \delta\}) < \epsilon$.

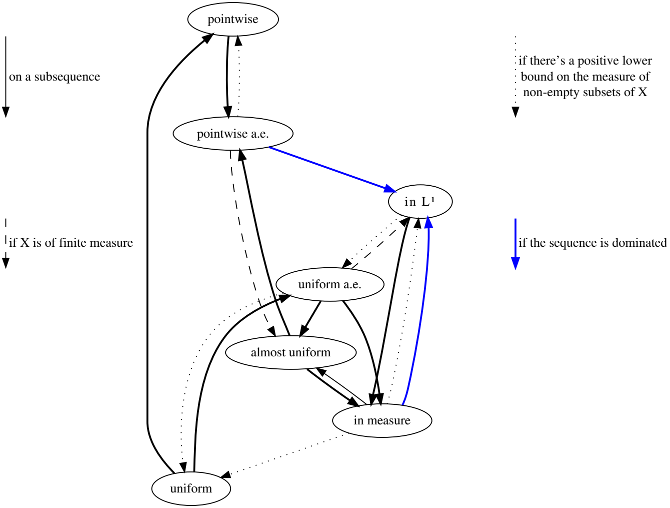
Locally uniform convergence, which we met in Section 1.3, isn't meaningful for general measure spaces. There are actually three versions of locally uniform convergence discussed there, referring to bounded subsets, open subsets or compact subsets. You would need a metric, or at least a topology, on X to make sense of those.

A comments on "norms" and uniqueness

- The "L^∞ norm" is not a norm, it's a semi-norm, just as with the "L^1 norm". There are also L^p norms for $1 < p < \infty$, but ignore those for now.
- You can make any of these a norm by quotienting by the subspace of functions f such that $\mu(\{x \in X : f(x) \neq 0\}) = 0$. In this module we don't take the quotient, but we still refer to it as and "the L^1 norm" and "the L^∞ norm".
- Limits in normed vector spaces are unique, but L^1 and L^∞ aren't really normed vector spaces. Limits in those "norms" are only unique up to changing the function on a set of measure zero.
- As long as we're willing to accept uniqueness up to changing the function on a set of measure zero, there's a much stronger uniqueness statement (Proposition 1.5.7): If $f_n \to g$ in any one of these seven senses and $f_n \to h$ in any of the other six then g and h agree almost everywhere.

Relations between modes of convergence

You might find this diagram helpful:



Explanation of the diagram

There are seven types of convergence in the diagram, with definitions given above. There are five types of arrows, indicating that one type of convergence implies another with or without additional hypotheses.

- Bold black: The first type of convergence implies the second with no additional hypotheses, e.g. if a sequence of functions converges almost uniformly then it converges pointwise almost everywhere.
- Solid black: The first type needn't imply the second, but you can always find a subsequence with the second type of convergence, e.g. if a sequence converges in measure then it has a subsequence which converges almost uniformly.

Explanation of the diagram, continued

- Dashed black: If $\mu(X) < \infty$ then the first type of convergence implies the second, e.g. if a sequence of functions on a set of finite measure converges almost everywhere then it converges in L^1. This applies in particular to probability spaces, which I'll discuss soon.
- Dotted black: If there is a $\kappa > 0$ such that $E \neq \emptyset \Rightarrow \mu(E) \geq \kappa$ then the first type of convergence implies the second, e.g under these conditions implies uniform convergence almost everywhere, which in turn implies uniform convergence. The condition looks strange, but it holds for counting measure with $\kappa = 1$.
- Solid blue: If there's an absolutely integrable G such that $|f_n| \leq G$ then the first type of convergence implies the second, e.g. pointwise convergence or convergence in measure implies convergence in L^1.

Probability interpretation of Measure Theory

Section 2.3 is about applying Measure Theory to Probability Theory. We won't discuss it in any detail, but it gives some insight into the modes of convergence.

A measure space $(X, \mathcal{B}, \mu)$ is called a probability space if $\mu(X) = 1$. X is interpreted as the sample space, roughly the set of all things which could happen. $\mathcal{B}$ is interpreted as the event space, roughly the set of all statements to which we assign probabilities. $\mu(E)$ is the probability that the event E occurs. μ-measurable functions on X are interpreted as random variables, roughly numbers which depend on what happens in such a way that it's meaningful to ask with what probability they lie in a given open set. The integral of such a function is interpreted as the expected value of the variable. This is "expectation" in the sense of probability theory, not a prediction.

If the probability that an event does not occur is zero then we say that it occurs almost surely.

Probability interpretation of modes of convergence

In what senses can a sequence of random variables tend to a random variable? In the probability interpretation, $x \in X$ is something which could happen and $f(x)$ is the value the corresponding random variable will take if x does indeed happen. So statements about convergence of functions correspond to statements about sequences of random variables:

- $f_n \to g$ pointwise: No matter what happens, the value of f_n will tend to that of g.
- $f_n \to g$ uniformly: No matter what happens, the value of f_n will tend to that of g, and the rate at which this happens can be quantified without knowing what will happen.
- $f_n \to g$ pointwise almost everywhere: Almost surely the value of f_n will tend to that of g.

Probability interpretation of Measure Theory

- $f_n \to g$ uniformly almost everywhere: Almost surely, the value of f will tend to that of g, at a rate independent of what happens.
- $f_n \to g$ almost uniformly: For every $\delta > 0$ there is an event with probability at most δ with the property that if it doesn't occur then the value of f_n tends to that of g with a rate which may depend on δ but doesn't depend on what happens.
- $f_n \to g$ in L^1 norm: The expected absolute difference between f_n and g tends to zero as n tends to infinity.
- $f_n \to g$ in measure: For every $\epsilon > 0$ the probability that the absolute difference between the value of f_n and that of g is at least ϵ tends to zero as n tends to infinity.