Formalising Mathematics

Project 3

[redacted]

March 26, 2022

Introduction

For my final project in *Formalising Mathematics*, I decided to built on top of my work from the second project, and formalise a notion in measure theory known as uniform integrability. Uniform integrability is a concept found both in measure theory as well as in probability theory. In this project, I will formalise their definitions and prove a powerful result known as the *Vitali convergence theorem*.

The Vitali convergence theorem generalises the dominated convergence theorem in the finite measure case and relates convergence in L^p to convergence in probability. This result is applicable in many situations where we are required to establish convergence in L^p and is fundamental in the martingale convergence theorem.

The content of this project has been merged into Lean's maths library - `mathlib` and can be found here. With this PR, I believe all measure-theoretic prerequisites for the formalisation of martingales are met and we can begin to prove non-trivial results about martingales.

Measure Theory vs. Probability Theory

It turns out there are two different definitions of uniform integrability for measure theorists and probability theorists. This project mostly focuses on the measure theorist's definition as it is the weaker of the two definitions while providing some basic APIs for the probability definition.

Definition (Measure Theorist's Uniform Integrability)**.** Let $(\mathcal{X}, \mathcal{F}, \mu)$ be a measure space. Then, a set of functions $S \subseteq L^1(\mu)$ is said to be uniformly integrable if for all $\epsilon > 0$, there exists some $\delta > 0$ such that

$$\|\mathbf{1}_A(f)\|_1 = \int_A |f| \mathrm{d}\mu < \epsilon$$

for all $f \in S$, $A \in \mathcal{F}$, $\mu(A) < \delta$.

Definition (Probability Theorist's Uniform Integrability)**.** Let $(\mathcal{X}, \mathcal{F}, \mu)$ be a measure space. Then, a set of functions $S \subseteq L^1(\mu)$ is said to be uniformly integrable if for all $\epsilon > 0$, there exists some K such that

$$\int_{\{|f| \geq K\}} |f| \mathrm{d}\mu < \epsilon$$

for all $f \in S$.

Straight away, we see that the probability definition implies the measure theory definitions as, for all $s \in \mathcal{F}$, we can decompose s by $s = (s \cap \{|f| \geq K\}) \cup (s \cap \{|f| < K\})$. Thus, integrating over the first set is bounded by ϵ and integrating over the second is limited by $K\mu(s)$. Thus, by taking $\mu(s) < K^{-1}\epsilon$, we have bounded the integral with 2ϵ.

Hence, with this in mind, one have the equivalent formulation of the probability definition of uniform integrability.

Proposition. A set of functions S is uniformly integrable in the probability sense if and only if it is uniformly integrable in the measure theory sense and is uniformly bounded in L^1.

As the second formulation incorporates the measure theory definition, it is easier to use the results and APIs for the measure theory definition in the probability case. Hence, it was decided to formalise the equivalent formulation of uniform integrability.

```
/- Uniform integrability in the measure theory sense. -/
def unif_integrable {m : measurable_space α} (f : ι → α → β)
  (p : ℝ≥0∞) (μ : measure α) : Prop :=
∀ ⦃ε : ℝ⦄ (hε : 0 < ε), ∃ (δ : ℝ) (hδ : 0 < δ),
  ∀ i s, measurable_set s → μ s ≤ ennreal.of_real δ →
  snorm (s.indicator (f i)) p μ ≤ ennreal.of_real ε

/- Uniform integrability in the probability theory sense. -/
def uniform_integrable {m : measurable_space α} [measurable_space β]
  (f : ι → α → β) (p : ℝ≥0∞) (μ : measure α) : Prop :=
(∀ i, measurable (f i)) ∧ unif_integrable f p μ ∧
  ∃ C : ℝ≥0, ∀ i, snorm (f i) p μ ≤ C
```

We note that our definition in Lean has the extra parameter `p`. Indeed, by again looking at our definition, uniform integrability requires the L^1 norm to be uniformly bounded. This is rather an arbitrary choice and we can in general consider the L^p norm for any $p \in [0, \infty]$ where we recover the traditional definition if $p = 1$.

Furthermore, as mentioned above, the Lean formulation of the probability definition of uniform integrability is equivalent to the standard definition found in literature. This equivalence is formalised in this project as follows.

```
lemma uniform_integrable_iff [is_finite_measure μ]
  (hp : 1 ≤ p) (hp' : p ≠ ∞) :
  uniform_integrable f p μ ↔ (∀ i, measurable (f i)) ∧
  ∀ ε : ℝ, 0 < ε → ∃ C : ℝ≥0,
  ∀ i, snorm ({x | C ≤ ‖f i x‖₊}.indicator (f i)) p μ ≤ ennreal.of_real ε
```

Some useful properties can be deduced straight away from these definitions.

Proposition. A single L^1 function is uniformly integrable.

Proof. We will prove this for the probability definition which will imply the measure theoretic definition. Namely, for $f \in L^1$, we will show that for all $\epsilon > 0$, there exists some C such that

$\int_{\{|f|\geq C\}} |f| \mathrm{d}\mu \leq \epsilon$, or equivalently,

$$\lim_{C\to\infty} \int_{\{|f|\geq C\}} |f| \mathrm{d}\mu = 0.$$

Indeed, we have

$$\lim_{C\to\infty} \int_{\{|f|\geq C\}} |f| \mathrm{d}\mu = \|f\|_1 - \lim_{C\to\infty} \int |f| \mathbf{1}_{\{|f|<C\}} \mathrm{d}\mu.$$

Thus, as $|f|\mathbf{1}_{\{|f|<C\}} \uparrow |f|$, by the monotone convergence theorem (or by the dominated convergence theorem as used in the formalisation), $\lim_{C\to\infty} \int |f|\mathbf{1}_{\{|f|<C\}} \mathrm{d}\mu = \|f\|_1$ and hence

$$\lim_{C\to\infty} \int_{\{|f|\geq C\}} |f| \mathrm{d}\mu = 0$$

as required. □

This is formalised in Lean as the following lemma and again, we note that the proof is generalised to L^p for all $1 \leq p < \infty$.

```
lemma uniform_integrable_subsingleton [subsingleton ι]
  (hp_one : 1 ≤ p) (hp_top : p ≠ ∞)
  (hf : ∀ i, measurable (f i)) (hf' : ∀ i, mem_ℒp (f i) p μ) :
  uniform_integrable f p μ
```

Corollary. A finite set of L^p functions is uniformly integrable.

Proof. Exercise. □

Vitali Convergence Theorem

One of the main results relying on the definition of uniform integrability is the Vitali convergence theorem.

Theorem (Vitali Convergence Theorem)**.** Let $(\mathcal{X}, \mathcal{F}, \mu)$ be a *finite* measure space. Then, a sequence of $L^1(\mu)$ functions $(f_n)_{n\in\mathbb{N}}$ converges in L^1 to some $f \in L^1$ if and only if $f_n \to f$ in measure and (f_n) is uniformly integrable.

Recalling the relations between the different notions of convergence in measure theory, in particular, we have

$$\text{convergence in } L^p\text{, convergence a.e.} \implies \text{convergence in } \mu$$

and

$$\text{convergence in } \mu \implies \exists \text{ subsequence converges a.e.}$$

Thus, with the Vitali convergence theorem, in some sense we complete the relation between the three as convergence in any one of the three will imply certain results about the other two.

Proof. ($\implies$) During the last project, we showed that convergence in L^1 implies convergence in measure and so, it suffices to show that (f_n) is uniformly integrable.

For all $\epsilon > 0$, choose $n_0 \in \mathbb{N}$ such that

$$\int |f_n - f| \mathrm{d}\mu < \frac{\epsilon}{2}$$

for all $n \geq n_0$. Then, as $\{f, f_1, \cdots, f_{n_0}\}$ is finite, it is uniformly integrable. Hence, there exists some $\delta > 0$ such that for all $A \in \mathcal{F}$ with $\mu(A) < \delta$,

$$\int_A |f| \mathrm{d}\mu < \frac{\epsilon}{2}, \text{ and } \max_{n \leq n_0} \int_A |f_n| \mathrm{d}\mu < \frac{\epsilon}{2}.$$

Now, choosing this to be our δ, for $n > n_0$,

$$\int_A |f_n| \mathrm{d}\mu = \int_A |f_n - f + f| \mathrm{d}\mu \leq \int_A |f_n - f| \mathrm{d}\mu + \int_A |f| \mathrm{d}\mu < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

and we have (f_n) is uniformly integrable as required.

($\impliedby$) Suppose for contradiction $\limsup_{n \to \infty} \int |f_n - f| \mathrm{d}\mu > 0$. Then, there exists some subsequence $\Lambda \subseteq \mathbb{N}$, such that

$$\lim_{n \in \Lambda \to \infty} \int |f_n - f| \mathrm{d}\mu > 0,$$

and $f_n \to f$ a.e. as $n \in \Lambda \to \infty$ (since convergence in measure implies the existence of such a subsequence). Now, as (f_n) is uniformly integrable, so is $(f_n) \cup \{f\}$. Thus, for all $\epsilon > 0$, there exists some $\delta > 0$ such that for all $A \in \mathcal{F}$, $\mu(A) < \delta$, we have

$$\int_A |f| \mathrm{d}\mu < \frac{\epsilon}{3}, \text{ and } \int_A |f_n| \mathrm{d}\mu < \frac{\epsilon}{3}$$

for all $n \in \mathbb{N}$. Applying Ergorov's theorem to $(f_n)_{n \in \Lambda}$, one finds a measurable set F such that $\mu(\mathcal{X} \setminus F) < \delta$ and

$$\sup_{x \in F} |f_n(x) - f(x)| \to 0,$$

as $n \in \Lambda \to \infty$. Thus, choosing n_0 such that

$$\sup_{x \in F} |f_n(x) - f(x)| < \frac{\epsilon}{3\mu(\mathcal{X})},$$

for all $n \in \lambda_{\geq n_0}$, it follows that for all $n \in \lambda_{\geq n_0}$,

$$\int |f_n - f| \mathrm{d}\mu = \int |f_n - f| \mathbf{1}_F \mathrm{d}\mu + \int |f_n - f| \mathbf{1}_{\mathcal{X} \setminus F} \mathrm{d}\mu \leq \frac{\epsilon}{3\mu(\mathcal{X})} \mu(F) + \frac{\epsilon}{3} + \frac{\epsilon}{3} \leq \epsilon,$$

which contradicts our assumption. $\square$

This is formalised in Lean as the following lemma.

```
lemma tendsto_in_measure_iff_tendsto_Lp
  [second_countable_topology β] [is_finite_measure μ]
  (hp : 1 ≤ p) (hp' : p ≠ ∞)
  (hf : ∀ n, mem_ℒp (f n) p μ) (hg : mem_ℒp g p μ) :
  tendsto_in_measure μ f at_top g ∧ unif_integrable f p μ ↔
  tendsto (λ n, snorm (f n - g) p μ) at_top (𝓝 0)
```

We note that the Lean formalisation in fact establishes the equivalence for L^p for all $1 \leq p < \infty$. Indeed, by looking at the proof, we note that we can generalise the lemma for arbitrary L^p.

Comment on the Proof Process

The formalisation of the Vitali convergence theorem was rather tedious. As we are working with three different types though out the formalisation, namely ℝ, `nnreal` and `ennreal`, coercions and maps between them are plentiful. An example of this can be seen in the very definition of uniform integrability itself. Indeed, as the type of ϵ is ℝ, to compare it with a measure which has type `ennreal`, we are forced to map ϵ using `ennreal.of_real`. This process is somewhat unavoidable by the definition of the L^p-norm though perhaps one may experiment with using different choices in the definition of uniform integrability (indeed, Remy has suggested to change ϵ to `ennreal`). As a result of this, large part of this project is simply computation which is mostly omitted from the proof on paper.

Generalized Dominated Convergence

The Vitali convergence theorem is often known as the generalized dominated convergence theorem. While the there is a version of Vitali convergence theorem which applies to infinite measure spaces, the one we have considered in this project only covers finite measure space. Nonetheless, in this case, the we may obtain a version of the dominated convergence theorem from the Vitali convergence theorem.

Lemma. A sequence of uniformly convergent L^1 functions is uniformly integrable.

Proof. Take N such that for all $n \geq N, |f_n - f| < 1$. Then, for all $\epsilon > 0$, take $\delta > 0$ such that for all sets s, $\mu(s) < \delta$, $\int_s |f| \mathrm{d}\mu < \epsilon$. Then, as

$$\int_s |f_n - f| \mathrm{d}\mu \leq \int_s 1 \mathrm{d}\mu = \mu(s) < \delta,$$

we have

$$\int_s |f_n| \mathrm{d}\mu \leq \int_s |f| + |f_n - f| \mathrm{d}\mu < \epsilon + \delta.$$

Hence, taking $\delta < \epsilon$ finishes the proof as there is only finitely many f_n with $n \leq N$. □

Recalling the Egorov's theorem which states that a sequence of almost everywhere convergent sequence, converges uniformly except on a set of arbitrary small measure. If $f_n \to f$

a.e. For all $\epsilon > 0$, take $f_n \to f$ uniformly on s^c where $\mu(s) < \epsilon/2$. By the above lemma, f_n is uniformly integrable on s^c and so by Vitali,

$$\int_{s^c} |f_n - f| \mathrm{d}\mu \to 0.$$

On the other hand, if $|f_n| \leq g$ a.e. (which is the condition required for dominated convergence), we have

$$\int_s |f_n - f| \mathrm{d}\mu \leq \int_s g + |f| \mathrm{d}\mu \to 0$$

as $\mu(s) \to 0$ as a single function is uniformly integrable.

While this is result is not proved in this project (as we already have dominated convergence theorem in `mathlib`), it is straight forward and could be a fun result to prove as the second project for the next cohort of students of *Formalising Mathematics*.

Future Work

As mentioned in the introduction, the Vitali convergence theorem is an important prerequisite for the martingale convergence theorem and is very useful in general to establish convergence in L^p (as in some way, convergence in measure is much weaker than convergence in L^p).

In the near future, we will build on this work to establish nontrivial results in martingale theory.

Another interesting result that builds on top of this is the Dunford-Pettis theorem for L^1 functions.

Theorem (Dunford-Pettis Theorem)**.** A sequence (f_n) of L^1 functions is uniformly integrable if and only if it is weakly compact in the weak topology on L^1.

We note the connection of Dunford-Pettis with Prokhorov's theorem.

Definition (Tight)**.** A family of probability measures $\mathcal{P}$ is said to be tight if for all $\epsilon > 0$, there exists a compact set K such that for all $\mathbb{P} \in \mathcal{P}$, $\mathbb{P}(K^c) < \epsilon$.

Theorem (Prokhorov)**.** A family of probability measures is tight if and only if it is relatively compact in the weak topology of measures.

This theorem can be generalised to finite signed measures for which we note an immediate connection to the Dunford-Pettis theorem. Indeed, by recalling that for all $f \in L^1$, $f\mu$ is a finite signed measure. Hence, applying Prokhorov's theorem on $(f_n\mu)$, we have $(f_n\mu)$ is tight if and only if it is relatively compact. Thus, by observing $(f_n\mu)$ is relatively compact if and only if (f_n) is relatively compact in the weak topology of L^1, and the tightness of $(f_n\mu)$ closely relates to the uniform integrability of (f_n), we note that Dunford-Pettis closely relates to Prokhorov as claimed.

On the other hand, one may obtain the Prokhorov theorem from Dunford-Pettis by a simple application of the Riesz-Markov-Kakutani representation theorem.

The actual proof of Dunford-Pettis is more elementary and is very doable in Lean. In particular, the main prerequisite of the Banach-Alaogu theorem is already in `mathlib`. I believe this can be a challenging third project for students of *Formalising Mathematics* in the future.