Advanced Real Analysis Tutorial 02

Yu Junao

October 13, 2024

Contents

1 Review

1.1 Measurable Functions

When it comes to analysis, we always focus on functions. After establishing the measure theory, we are supposed to applies it to functions.

There are a lot discontinuous functions like characteristic functions which still matter. We need to expand the varieties of functions.

Definition 1 (Measurable function)**.** *Let* $(X, \mathcal{M}, \mu)$ *and* $(Y, \mathcal{N}, \nu)$ *be two measure spaces. A function* $f : X \to Y$ *is measurable if*

$$f^{-1}(E) \in \mathcal{A}, \ \forall E \in \mathcal{B}.$$

In other words, f^{-1} brings a measurable set to a measurable set (the inverse proposition is not necessarily correct).

Particularly taking $\mathcal{N} = \mathcal{B}_Y$, we obtain an equivalent definition that f^{-1} brings an open set to a measurable set. Additionally, we have the following theorem since open sets belong to the Borel σ-algebra.

Theorem 1 (Compatibility between continuity and measurability)**.** *A continuous function mapping* $(X, \mathcal{B}_X, \mu)$ *into* $(Y, \mathcal{B}_Y, \nu)$ *is measurable (**Borel measurable**).*

Here we require topological structures on X and Y.

Remark 1. *In practice, we are more concerned about **Lebesgue measurable** functions, whose preimage of a Borel set is a Lebesgue set. The composition of two Lebesgue measurable functions is not necessarily a Lebesgue measurable function.*

Measurable functions are very "flexible". The collection of measurable functions if closed under linear combinations, multiplications, and even limits.

The most common measurable functions are **simple functions**.

Definition 2 (Simple function)**.** *A simple function is a finite linear combination of characteristic functions of measurable sets,*

$$f = \sum_{j=1}^{n} a_j \chi_{A_j}.$$

Remark 2. *Finiteness is required for simple functions. As a result, the floor function* $g(x) = [x]$ *is not a simple function.*

According to Taylor expansion, an analytic functions is approximated by polynomials, which are somehow simpler analytic functions. Measurable functions possess a similar property.

Theorem 2 (Approximation of simple functions)**.** *A measurable function* f *could be pointwise approximated by a increasing sequences of simple functions. In particular if* f *is bounded, then the approximation is uniform.*

1.2 Lebesgue Integration

With the help of the approximation of simple functions, Lebesgue integration theory is established based on measure theory.

Definition 3 (Lebesgue integral of non-negative measurable functions)**.** *Let*

$$\varphi(x) = \sum_{j=1}^{n} a_j \chi_{A_j}$$

be a non-negative simple function on $(X, \mathcal{M}, \mu)$*, then we define*

$$\int_X \varphi \, d\mu = \sum_{j=1}^{n} a_j \mu(A_j).$$

Additionally for a non-negative measurable function f*, define*

$$\int_X f \, d\mu = \sup \left\{ \int_X \varphi \, d\mu \middle| \varphi \leq f, \varphi \text{ is simple} \right\}.$$

We can also define the integration on an arbitrary measurable set by

$$\int_E f \, d\mu = \int_X f \chi_E \, d\mu.$$

For general non-negative measurable function $f = f^+ - f^-$, we can define the integral of its positive and negative parts respectively.

Definition 4 (Integrability)**.** *We can define the integral of* f *if*

$$\int f^+ < +\infty \text{ or } \int f^- < +\infty.$$

And f *is integrable if and only if*

$$\int |f| = \int f^+ + \int f^- < +\infty.$$

We denote $L^1(X, \mu) = L^1(X) = L^1(\mu) = L^1$ *as the collection pf all integrable functions on* X *with respect of measure* μ*.*

In fact, L^1 is a **Banach space** (complete normed linear space). The value of an integration on a null set is zero, thus two almost everywhere equal functions are the same in L^1.

The most remarkable highlight of Lebesgue integration lies in the convenience of "swapping integrations". The following three theorems are equivalent.

Theorem 3 (Monotone convergence theorem)**.** *For an increasing sequence of non-negative measurable functions* $\{f_j\}_{j=1}^{\infty}$ *that converges to* f *almost everywhere, we have*

$$\lim_{j\to\infty}\int f_j = \int \lim_{j\to\infty} f_j = \int f.$$

Theorem 4 (Fatou's lemma)**.** *For a sequence of measurable functions* $\{f_j\}_{j=1}^{\infty}$*, we have*

$$\int \varliminf_{j\to\infty} f_j \le \varliminf_{j\to\infty} \int f_j.$$

Theorem 5 (Dominant convergence theorem)**.** *For a sequence of functions* $\{f_j\}_{j=1}^{\infty}$ *that converges to* f *almost everywhere, if* $|f_j| \le g \in L^1$*, then we have*

$$\lim_{j\to\infty}\int f_j = \int \lim_{j\to\infty} f_j = \int f.$$

Remark 3. *The strict inequality in Fatou's Lemma is achieved by*

$$f_j = \chi_{[j,j+1]} \text{ or } g_n = n\chi_{[0,\frac{1}{n}]}.$$

This fact implies the importance of dominant functions.

Remark 4. *The condition* $|f_j| \le g \in L^1$ *could be weaken to*

$$|f_j| \le g_j \in L^1,\ g_j \to g,\ \text{almost everywhere},$$

which appeared in our homework.

We have investigated into the criteria of swapping a limit and an integration. To swap two integrations, we must comprehend the concepts of repeated integration and multiple integration. Therefore, we need the help of product measures.

When defining the product of two objects, we always consider their generators for a start.

Definition 5 (Product measure space)**.** *Let* $(X, \mathcal{M}, \mu)$ *and* $(Y, \mathcal{N}, \nu)$ *be two measure spaces. We can define product measure space* $(X \times Y, \mathcal{M} \times \mathcal{N}, \mu \times \nu)$ *where* $\mathcal{M} \times \mathcal{N}$ *is the* σ*-algebra generated by*

$$\{A \times B \mid A \in \mathcal{M}, B \in \mathcal{N}\},$$

and $\mu \times \nu(A \times B) = \mu(A) \times \nu(B)$*.*

Remark 5. *Of course, we can make the product measure space complete and define the Lebesgue σ-algebra on $X \times N$.*

Recall the definition of the sections of a set measurable set or a measurable function, and we have the desired theorem in the sense of measures.

Theorem 6. *Let $E \in \mathcal{M} \times \mathcal{N}$, then the two functions $x \to E_x$ and $y \to E^y$ are both measurable. Additionally, we have*

$$\mu \times \nu(E) = \int \nu(E_x)\, d\mu(x) = \int \mu(E^y)\, d\nu.$$

Based on last theorem, we reach "the peak of real analysis".

Theorem 7 (Fubini-Tonelli theorem)**.** *If $f \in L^+(X \times Y) \cup L^1(X \times Y)$, then*

$$\int f\, d(\mu \times \nu) = \int \left(\int f(x,y)\, d\nu(y) \right) d\mu(x) = \int \left(\int f(x,y)\, d\mu(x) \right) d\nu(y).$$

1.3 Signed Measures

With the help of measure, integration theory is established. Conversely, we can construct a "measure" through

$$\mu_f(E) = \int_E f\, \mathrm{d}\mu,\ f \in L^1(\mu).$$

Here we call μ_f is **absolutely continuous** with respect of μ since $\mu(E) = 0$ implies $\mu_f(E) = 0$, denoted as $\mu_f \ll \mu$. It is easy to verify that μ_f is definitely a measure when $f \in L^+(\mu)$. But once f takes negative value, μ_f is no longer a measure since we require measures take only non-negative values. However, it is not just nonsense.

Definition 6 (Signed measure)**.** *A set function $\nu : \mathcal{M} \to [-\infty, +\infty]$ is called a signed measure if it satisfies the two conditions of measure yet not achieves $\pm\infty$ simultaneously.*

Remark 6. *Signed measures are not measures, as manifolds with boundary are not manifolds. To aviod the situation $(+\infty) - (+\infty)$, signed measures take values at most onf of $+\infty$ and $-\infty$.*

Since functions can be decomposed into positive and negative part, we hope the same is true for measures.

Theorem 8 (Hahn decomposition)**.** *For a signed measure ν on measure space $(X, \mathcal{M})$, there are P, Q such that $X = P \sqcup Q$ and ν is positive on P while negative on Q. Such composition is unique in the sense of measure.*

Theorem 9 (Jordan decomposition)**.** *For a signed measure ν on measure space $(X, \mathcal{M})$, there are two measures ν^+ and ν^- such that $\nu = \nu^+ - \nu^-$ and $\nu^+ \perp \nu^-$.*

Remark 7. *Practically, we usually apply only one of them.*

Here $\nu^+ \perp \nu^-$ means ν^+ and ν^- are **singular**, they never take nonzero values on the same set.

Finally, we state the most significant theorem in signed measure.

Theorem 10 (Radon-Nikodym theorem)**.** *Given a σ-finite measure μ and a σ-finite signed measure ν, we could decompose ν into $\rho + \lambda$, where*

$$\begin{aligned} &d\rho = f\, d\mu, \ \ f^+ \in L^1 \ \text{or} \ f^- \in L^1, \\ &\lambda \perp \mu. \end{aligned}$$

Particularly, $f = \frac{d\rho}{d\mu}$ is called the ***Radon-Nikodym derivative****.*

This theorem indicates the fact that we can separate a "relatively good" measure form ν such that the remaining part "takes no effect" on the domain.

1.4 Foundations of L^p Spaces

Inspired by the p-norms in Euclidean spaces, we develop L^p norms from L^1 norm.

Definition 7 (L^p space)**.** *Consider the norm*

$$\begin{aligned} \|f\|_p &= \|f\|_{L^p} = \left(\int |f|^p \right)^{\frac{1}{p}}, 1 \le p < +\infty, \\ \|f\|_\infty &= \|f\|_{L^\infty} = \inf_Z \left\{ \sup_{x \in X \setminus Z} |f| \,\middle|\, \mu(Z) = 0 \right\}. \end{aligned}$$

We define L^p space as the collection of all functions whose L^p norm is finite. In fact, it is a Banach space.

An important issue in L^p space is inequalities. Step by step, we proved three inequalities.

Theorem 11 (Young's inequality)**.** *For $a, b \ge 0$, we have*

$$ab \le \frac{a^p}{p} + \frac{b^{p'}}{p'}.$$

The equality holds if and only if $a^p = b^{p'}$.

Theorem 12 (Hölder's inequality)**.** *For $f \in L^p$ and $g \in L^{p'}$, we have*

$$\|fg\|_1 \leq \|f\|_p \|g\|_{p'}.$$

The equality holds if and only if $|f|^p$ and $|g|^{p'}$ are linearly dependent in L^1.

This inequality implies L^p and $L^{p'}$ are dual spaces.

Theorem 13 (Minkovski's inequality)**.** *For $f, g \in L^p$ and $g \in L^{p'}$, we have*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

The equality holds if and only if f and g are linearly dependent in L^p.

This inequality implies $\|\cdot\|_p$ is a norm for $1 \leq p \leq +\infty$.

Remark 8. *There is a counterpart of Hölder's inequality for multiple functions, which is*

$$\|f_1 \cdots f_n\|_p \leq \|f_1\|_{p_1} \cdots \|f_n\|_{p_n}, \ \frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_n}.$$

An interesting corollary is the following theorem.

Theorem 14 (Interpolation inequality)**.** *For $1 \leq p < q < r \leq +\infty$, we have $L^p \cap L^r \subset L^q$, and*

$$\|f\|_q \leq \|f\|_p^{\lambda} \|f\|_r^{1-\lambda}, \ \frac{1}{q} = \frac{\lambda}{p} + \frac{1-\lambda}{r}.$$

In fact, we have another conclusion for such p, q, r, which is $L^q \subset L^p + L^r$. To comprehend this phenomenon clearly, we decompose a L^q function f into

$$f = f\chi_{|f| \leq 1} + f\chi_{|f|>1}.$$

It is easy to check the first term belongs to L^r while the second term belongs to L^p.

2 Solutions to Homework

2.1 Exercise 1.5.25

Proof. We only need to consider the case $\mu(E) = +\infty$. Let

$$E_n = [n, n+1) \cap E, \ n \in \mathbb{Z}$$

thus E_n is a measurable and $\mu(E_n) \leq 1 < +\infty$. By previous conclusion, for any $k > 0$, there exists a F_σ set $K_{k,n} \subset E_n$ such that

$$\mu(E_n \backslash K_{k,n}) \leq \frac{1}{2^{|n|+k}}$$

Take union, we obtain $K_k \subset E$ for F_σ set

$$K_k = \bigcup_{n=-\infty}^{\infty} K_{k,n}$$

and

$$\mu(E \backslash K_k) \leq \frac{3}{2^k}.$$

Therefore, the F_σ set

$$K = \bigcup_{k=1}^{\infty} K_n \subset E$$

satisfies $\mu(E \backslash K) = 0$. We have proved a measurable set can be interiorly approximated by an F_σ set.

Applying this conclusion, we construct an F_σ set K such that $K \subset E^c$ and $\mu(E^c \backslash K) = 0$. Hence $U = K^c$ is a G_δ set including E, such that $\mu(U \backslash E) = 0$. $\square$

Remark 9. *It is also accessible to construct a G_δ set U that covers E directly. However, we need to construct a family of open sets $\{U_{n,k}\}_{n,k=1}^{\infty}$ instead of G_δ sets, or U will become $G_{\delta\sigma}$.*

2.2 Exercise 2.1.2

(1)

Proof. First, we need a conclusion that $f : X \to \overline{\mathbb{R}}$ is measurable if f is measurable on $f^{-1}(\mathbb{R})$ and $f^{-1}(\pm\infty) \in \mathcal{M}$.

For $E \in \mathcal{B}_{\overline{\mathbb{R}}}$, $f^{-1}(E)$ is measurable if $E \in \mathcal{M}_{\mathbb{R}}$. Otherwise, we without loss of generality assume $E = A \cup \{+\infty\}$ where $A \in \mathcal{B}_{\mathbb{R}}$, then

$$f^{-1}(E) = f^{-1}(A) \cup f^{-1}\{+\infty\} \in \mathcal{M}.$$

Back to the original problem, we have

$$\begin{aligned}(fg)^{-1}\{+\infty\} = &\left(f^{-1}\{+\infty\} \cap g^{-1}(0, +\infty]\right) \cup \left(f^{-1}(0, +\infty] \cap g^{-1}\{+\infty\}\right) \\ &\cup \left(f^{-1}\{-\infty\} \cap g^{-1}[-\infty, 0)\right) \cup \left(f^{-1}[-\infty, 0) \cap g^{-1}\{-\infty\}\right) \in \mathcal{M}.\end{aligned}$$

Similarly, we have $(fg)^{-1}\{-\infty\} \in \mathcal{M}$. As we know, fg is measurable on $(fg)^{-1}(\mathbb{R})$ since f and g are respectively measurable on $f^{-1}(\mathbb{R})$ and $g^{-1}(\mathbb{R})$. According to the conclusion above, fg is measurable on X. $\square$

(2)

Proof. If $a = \pm\infty$, we without loss of generality assume $a = +\infty$. In that case, $f + g$ is obviously measurable on $h^{-1}(\mathbb{R})$. On the other hand, we have

$$\begin{aligned} h^{-1}\{+\infty\} = &\big(f^{-1}\{+\infty\} \cap g^{-1}(\mathbb{R})\big) \cup \big(f^{-1}(\mathbb{R}) \cap g^{-1}\{+\infty\}\big) \\ &\cup \big(f^{-1}\{+\infty\} \cap g^{-1}\{-\infty\}\big) \cup \big(f^{-1}\{-\infty\} \cap g^{-1}\{+\infty\}\big) \in \mathcal{M}, \\ h^{-1}\{-\infty\} = &\big(f^{-1}\{-\infty\} \cap g^{-1}(\mathbb{R})\big) \cup \big(f^{-1}(\mathbb{R}) \cap g^{-1}\{-\infty\}\big) \in \mathcal{M}. \end{aligned}$$

If $a \in \mathbb{R}$, we can similarly show that $h^{-1}\{\pm\infty\} \in \mathcal{M}$. Let $E \in \mathbb{R}$ excluding a, then $h^{-1}(E)$ is obviously measurable. Otherwise, we have

$$\begin{aligned} h^{-1}(E) = &(f+g)^{-1}(E \backslash \{a\}) \cup (f+g)^{-1}\{a\} \\ &\cup \big(f^{-1}\{+\infty\} \cap g^{-1}\{-\infty\}\big) \cup \big(f^{-1}\{-\infty\} \cap g^{-1}\{+\infty\}\big) \in \mathcal{M}. \end{aligned}$$

In summary, h is always measurable. ☐

Remark 10. *Another elegant approach is to consider the explicit expression*

$$h(x) = a\chi_A + (f(x) + g(x))\chi_{A^c},$$

where

$$A = \big(f^{-1}(+\infty) \cap g^{-1}(-\infty)\big) \cup \big(f^{-1}(-\infty) \cap g^{-1}(+\infty)\big)$$

is measurable.

Remark 11. *The statement "f is measurable on $f^{-1}(\mathbb{R})$" is not equivalent with "$f^{-1}(\mathbb{R})$ is measurable".*

2.3 Exercise 2.5.49

Proof. For a null set $E \in \mathcal{M} \times \mathcal{N}$, we have

$$0 = \mu \times \nu(E) = \int \nu(E_x)\,\mathrm{d}\mu(x) = \int \mu(E^y)\,\mathrm{d}\nu(y),$$

which implies $\nu(E_x) = \mu(E^y) = 0$ for almost every x and y.

For the λ-null set

$$A = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) \neq 0\},$$

there exists an $E \in \mathcal{M} \times \mathcal{N}$ including E_0 such that $\mu \times \nu(E) = 0$. Therefore,

$$\int |f_x| \, \mathrm{d}\nu(y) = \int \chi_{A_x} |f_x| \, \mathrm{d}\nu(y) = 0,$$
$$\int |f^y| \, \mathrm{d}\mu(x) = \int \chi_{A^y} |f^y| \, \mathrm{d}\mu(x) = 0.$$

Back to the proof of the theorem 2.39, suppose $f \in L^+(\lambda) \cup L^1(\lambda)$. According to Proposition 2.12, there exists a $\mu \times \nu$-measurable function g such that $f = g$, λ-almost everywhere. Since g_x is ν-measurable and g^y is μ-measurable.

Define a function $h = f - g$ that equals 0 almost everywhere, then the latter lemma implies that h_x is ν-measurable for almost every x while h^y is μ-measurable for almost every y.

Particularly when $f \in L^1(\lambda)$, we have $h_x \in L^1(\nu)$ for almost every x. By Fubini's theorem, $g_x \in L^1(\nu)$ for almost every x, thus $f_x \in L^1(\nu)$ for almost every x. The counterpart for f^y is still correct. Moreover, we have almost everywhere

$$\int h_x \, \mathrm{d}\nu(y) = 0 \Longrightarrow \int g_x \, \mathrm{d}\nu(y) = \int f_x \, \mathrm{d}\nu(y),$$
$$\int h^y \, \mathrm{d}\mu(x) = 0 \Longrightarrow \int g^y \, \mathrm{d}\mu(x) = \int f^y \, \mathrm{d}\nu(x).$$

If $f \in L^+(\lambda)$, then $g \in L^+(\mu \times \nu)$, thus Tonelli's theorem indicates functions

$$x \to \int g_x \, \mathrm{d}\nu(y) = \int f_x \, \mathrm{d}\nu(y),$$
$$y \to \int g^y \, \mathrm{d}\mu(x) = \int f^y \, \mathrm{d}\nu(x)$$

are both measurable. Correspondingly, if $f \in L^1(\lambda)$, then $g \in L^1(\mu \times \nu)$, thus Fubini's theorem indicates the two functions above are both integrable. An ultimately application of Fubini-Tonelli's theorem leads to the identity

$$\int f \, \mathrm{d}\lambda = \int \left(\int f \, \mathrm{d}\mu \right) \mathrm{d}\nu = \int \left(\int f \, \mathrm{d}\nu \right) \mathrm{d}\mu.$$

□

Remark 12. *It is a sophisticated proof that few students managed to complete, though the theorem itself is far more significant than the details.*

2.4 Exercise 3.2.17

Proof. Consider a new measure ρ such that $\mathrm{d}\rho = f\,\mathrm{d}\mu$. For $E \in \mathcal{N}$, we have

$$\nu(E) = 0 \Longrightarrow \mu(E) = 0 \Longrightarrow \rho(E) = 0.$$

which implies $\rho \ll \nu$.

Let $g = \frac{\mathrm{d}\rho}{\mathrm{d}\nu}$ be the Radon-Nikodym derivative which is unique, and it is easy to check that

$$\int_E f\,\mathrm{d}\mu = \int_E g\,\mathrm{d}\nu.$$

$\square$

Remark 13. *This proposition is no more correct when μ is infinity. It is easy to check $\mathcal{M} = \mathcal{B}_{\mathbb{R}}, \mathcal{N} = \{0, \mathbb{R}\}, f = \chi_{[0,1]}$ is a counterexample. This phenomenon comes from the fact that the Radon-Nikodym derivative is not always integrable in an infinite measure space.*

3 Topics in Convergence

Convergence is always a popular issue in analysis. For example, we usually construct a sequence of "approximation solutions" in order to reach the real solution. If the sequence or one of its subsequences is convergent, the limit is usually what we desire.

3.1 Common Modes of Convergences

Literally, f_n converges to f means they are somehow going closer and closer. As we know, metric is used to describe the distant between two "vectors". Consequently, we can define convergence in metric by

$$d(f_n, f) \to 0,\ n \to \infty.$$

Since a norm induces a metric, we have

Definition 8 (Converge in norm)**.** *Let $\|\cdot\|$ be a norm on a linear space, then f_n converges to f in norm if*

$$\|f_n - f\| \to 0,\ n \to \infty.$$

It is sometimes called ***strong convergence****.*

As is learned in functional analysis, inner product induces norm, norm induces metric, and metric induces topology. We can actually extend the concept of convergence to general topological spaces.

Definition 9 (Converge in topology). *Let $(X, \mathcal{T})$ be a topological space, then x_n converges to x if $\forall\, U \in \mathcal{T}$, $\exists\, N$ such that*

$$x_n \in U,\ \forall\, n > N.$$

Besides, measure seems to be another concept relevant with "size".

Definition 10 (Converge in measure). *Let $(X, \mathcal{M}, \mu)$ be a measure space, then f_n converges to f in measure if*

$$\lim_{n\to\infty} \mu\{|f_n - f| > \varepsilon\} = 0,\ \forall\, \varepsilon > 0.$$

Remark 14. *Note that the ε here never tends to 0. We only take limit of n. Literally, convergence in measure means the set of points that diverge is not too large in the sense of measure.*

Recall what we learned in mathematical analysis, pointwise convergence and uniform convergence are the most common modes of convergence at that time. Fortunately, we can develop them to real analysis.

As null sets are indistinguishable under measures, we prefer the expression "almost everywhere".

Definition 11 (Almost everywhere convergence). *f_n converges to f almost everywhere if f_n converges to f except for a null set.*

Definition 12 (Almost uniform convergence). *f_n converges to f almost uniformly on E if $\forall\, \varepsilon > 0$, $\exists\, E_\varepsilon \subset E$ such that $\mu(E_\varepsilon) < \varepsilon$ and f_n converges to f uniformly on $E \backslash E_\varepsilon$.*

They are weaken than the counterparts in mathematical analysis. For example, x^n converges to 0 almost uniformly but not uniformly on $[0, 1]$.

When it comes to L^p spaces, we discussed its duality. Previously we mentioned strong convergence, and thus there is "weaker convergences".

Definition 13 (Weak convergence). *Let X be a normed linear space, then $f_n \in X$ converges to f weakly if*

$$Tf_n \to Tf,\ \forall\, T \in X^*.$$

*Here X^* denotes the **dual space** of X, which is composed of all bounded linear operators on X.*

Conversely,

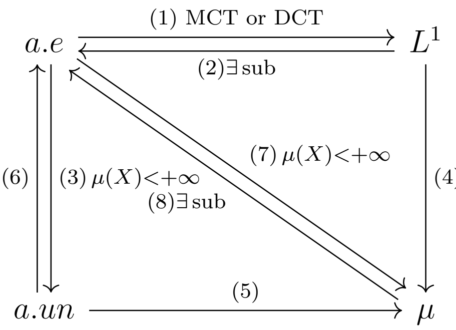
Definition 14 (Weak * convergence)**.** *Let X be a normed linear space, then $T_n \in X^*$ converges to T weakly * if*

$$T_n f \to Tf, \ \forall f \in X.$$

Remark 15. *Strong convergence implies weak convergence, and weak convergence implies weak * convergence. The inverse propositions are both incorrect.*

3.2 Relations

First, we consider the 4 elementary modes of convergence as following.



Theorem 15 (Relations among convergences)**.** *Relations of different modes of convergence are shown in the graph above.*

Proof. **(1):** Trivial.

(2): Attributed to (4) and (7).

(3):(Egorov's theorem) Let $f_n \to f$ almost everywhere on X and set

$$E_{k,n} = \bigcup_{j=n}^{\infty} \left\{ |f_j - f| \geq \frac{1}{k} \right\}.$$

For fixed k, it is obvious that $E_{k,n}$ decreases with respect to n and

$$\mu\left(\bigcap_{n=1}^{\infty} E_{n,k}\right) = \lim_{n\to\infty} \mu(E_{n,k}) = 0.$$

Given $\varepsilon > 0$, the continuity of measure implies the existence of N_k such that

$$\mu(E_{k,n}) < \frac{\varepsilon}{2^k}, \ \forall n > N.$$

As a result, we denote the union of $E_{k,n}$ with respect to n as E whose measure is no larger than ε, and f_n uniformly converges to f on $X\backslash E$ since

$$|f_n - f| < \frac{1}{k},\ \forall\, n > N_k.$$

(4):(Chebyshev's inequality) Direct computation shows

$$\mu(\{|f_n - f| \geq \varepsilon\}) = \frac{1}{\varepsilon}\int_{\{|f_n-f|\geq\varepsilon\}} 1 \leq \frac{1}{\varepsilon}\int_{\{|f_n-f|\geq\varepsilon\}} |f - f_n| \leq \frac{1}{\varepsilon}\int_X |f_n - f| \to 0,$$

as n tends to ∞ for any fixed $\varepsilon > 0$.

(5): $\forall\,\varepsilon > 0$, $\exists\, E$ such that $\mu(E) < \varepsilon$ and $f_n \rightrightarrows f$ on $X\backslash E$. In other words, for any $\delta > 0$, there exists an $N > 0$ such that for $n > N$ we have

$$|f_n(x) - f(x)| < \delta,\ \forall\, x \in X\backslash E.$$

As a result,

$$\{|f_n - f| \geq \varepsilon\} \subset E \Longrightarrow \mu\{|f_n - f| \geq \varepsilon\} < \delta.$$

That is to say, f_n converges to f in measure.

(6):$\forall\, k \geq 1$, $\exists\, E_k$ such that $\mu(X\backslash E_k^c) < \frac{1}{k}$ and $f_n \rightrightarrows f$ on $X\backslash E_k$. Consider

$$E = \bigcup_{k=1}^{\infty} E_k$$

such that $\mu(X\backslash E) = 0$, and it is easy to verify that $f_n \to f$ pointwise on E.

(7): Attributed to (3) and (5).

(8): Suppose f_n converges to f in measure. $\forall\, k > 0$, $\exists\, n_k > 0$, such that

$$\mu\left\{|f_n - f| \geq \frac{1}{k}\right\} < \frac{1}{2^k},\ \forall\, n \geq n_k$$

Define

$$E_k = \left\{|f_{n_k} - f| \geq \frac{1}{k}\right\} \text{ and } F_j = \bigcup_{k=j+1}^{\infty} E_k,$$

and we notice that $\mu(F_j) \leq \frac{1}{2^j}$. As a result, the intersection

$$Z = \bigcap_{j=1}^{\infty} F_j$$

is a μ-null set. As a result,

$$x \in X\backslash Z \Longrightarrow x \notin F_j, \exists\, j \Longrightarrow x \notin E_k, \forall\, k > j \Longrightarrow |f_{n_k}(x) - f(x)| < \frac{1}{k}, \forall\, k > j.$$

To conclude, f_{n_k} converges to f pointwise on $X\backslash Z$. □

Similar with L^1 convergence, we can define L^p convergence, which is much more useful in analysis. Actually, L^p convergence usually comes from inequalities, which a reason why we focus on inequalities in this chapter. For example, the corollary of Hölder's inequality for finite space

$$\|f\|_p \leq \mu(X)^{\frac{1}{p}-\frac{1}{q}}\|f\|_q,\ \forall\, q > p,$$

states the fact that L^q convergence implies L^p convergence on bounded domain.

In the future, we will learn topics in Sobolev spaces, where a variety of inequalities and embedding theorems will lead to different modes of convergence, playing an essential role in the existence and regularity problems of equations.

3.3 Application

We are going to introduce an application of different modes of convergence that was initiated by Brezis and Nirenberg in 1983.

Consider the following nonlinear partial differential equation on a bounded region Ω,

$$\begin{cases} -\Delta u = u^p + \lambda u, & x \in \Omega, \\ u > 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega. \end{cases}$$

where $p = \frac{n+2}{n-2}$. We shall investigate into its solvability in H_0^1 for $\lambda > 0$. Here

$$u \in H_0^1(\Omega) \Longleftrightarrow \begin{cases} u \in L^2(\Omega), \\ |\nabla u| \in L^2(\Omega), \\ u|_{\partial\Omega} = 0. \end{cases}$$

Consider a functional

$$\Phi(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 - \frac{1}{p+1}\int_\Omega |u|^{p-1} - \frac{\lambda}{2}\int_\Omega |u|^2.$$

A solution to the previous equation if and only if it is a **critical point** (extremal point) since calculus of variations implies

$$\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} \Phi(u + \varepsilon\varphi) = 0 \Longleftrightarrow -\Delta u = u^p + \lambda u$$

for $\forall\, \varphi \in C_c^\infty(\Omega)$.

Without loss of generality, we assume $\|u\|_{p+1} = 1$ since the critical points do not change under scaling. Define

$$S_\lambda = \inf_{\substack{u \in H_0^1 \\ \|u\|_{p+1}=1}} \left\{ \|\nabla u\|_2^2 - \lambda \|u\|_2^2 \right\}.$$

The H_0^1 norm is induced by the inner product

$$\langle u, v \rangle_{H_0^1} = \int uv + \int \nabla u \cdot \nabla v.$$

The equation is solvable if S_λ is achieved by some u.

Actually, it has been proved in Brezis-Nirenberg's paper that $S_\lambda < S_0$ strictly as $\lambda > 0$. and we have the following theorem.

Theorem 16 (Lieb)**.** *If $S_\lambda < S_0$, then S_λ is achieved.*

Proof. Let $\{u_j\}_{j=1}^\infty \subset H_0^1$ be a **minimizing sequence** of S_λ such that $\|u_j\|_{p+1} = 1$ and

$$\|\nabla u_j\|_2^2 - \lambda \|u_j\|_2^2 = S_\lambda + o(1).$$

Obviously, $\{u_j\}_{j=1}^\infty$ is bounded in H_0^1, thus the reflexivity extracts a weak convergent subsequence, still denoted as $\{u_j\}_{j=1}^\infty$. Additionally, we have

$$u_j \rightharpoonup u \text{ in } H_0^1 \Longrightarrow u_j \to u \text{ in } L^2 \Longrightarrow u_j \to u, \text{ a.e.}$$

As a result of Sobolev embedding and Riesz theorem.

Since $\{u_j\}_{j=1}^\infty$ is bounded in L^p, there is a subsequence that weakly converges to u in L^p. We need to prove $u_j \to u$ strongly in L^p.

Abstract $v_j = u_j - u$ such that

$$\begin{aligned} v_j &\rightharpoonup 0 \text{ in } H_0^1 \\ v_j &\to v, \text{ a.e.} \end{aligned}$$

Through a comprehensive application of **Sobolev inequality**, **Brezis-Lieb lemma**, and the definition of S_λ, we have

$$\begin{aligned} S_\lambda &= \int |\nabla u_j|^2 - \lambda \int u_j^2 + o(1) \\ &= \int |\nabla v_j|^2 - \lambda \int |v_j|^2 + \int |\nabla u|^2 - \lambda \int |u|^2 + o(1) \\ &\geq S_0 \|v_j\|_{p+1}^2 + S_\lambda \|u\|_{p+1}^2 + o(1) \\ &= (S_0 - S_\lambda)\|v_j\|_{p+1}^2 + S_\lambda \left(\|u_j\|_{p+1}^2 + \|u - u_j\|_{p+1}^2 \right) + o(1) \\ &= (S_0 - S_\lambda)\|v_j\|_{p+1}^2 + S_\lambda + o(1). \end{aligned}$$

Since $S_0 > S_\lambda$, we obtain $\|v_j\|_{p+1}^2 \to 0$. Moreover, we have

$$\|\nabla v_j\|_2^2 = \|\nabla u_j\|_2^2 - \|\nabla u\|_2^2 + o(1) = \lambda \|u_j\|_2^2 - \lambda \|u\|_2^2 + o(1) = o(1).$$

Therefore, $u_j \to u$ strongly in both L^{p+1} and H_0^1, and S_λ is achieved by this u. $\square$

This theorem implies the equation has a solution in H_0^1 for $\lambda > 0$. The solution is positive in Ω as a result of strong maximum principle.