APMA 2110: Real Analysis

Milan Capoor

Fall 2024

Chapter 1

Analysis and Metric Spaces

1.1 Sept 05

Some basic notation:

$$\begin{aligned}
\mathbb{N} &:= \{1, 2, 3, \ldots\} \\
\mathbb{Z} &:= \{\ldots, -2, -1, 0, 1, 2, \ldots\} \\
\mathbb{Q} &:= \left\{\frac{m}{n} \mid m, n \in \mathbb{Z}, n \neq 0\right\} \\
\mathbb{R} &:= \text{the set of real numbers} \\
\mathbb{C} &:= \text{the set of complex numbers}
\end{aligned}$$

Some basic logic:

- $(A \implies B) \iff (\neg B \implies \neg A)$ (contrapositive)
- $E \subset X \implies \forall x \in E,\ x \in X$

Sets

Note that in this course, $\subset$ includes the possibility of equality, while $\subsetneq$ does not.

Power Set: $P(X) = \{E : E \subseteq X\}$

Example: $X = \{1, 2, 3\}$

$$P(X) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}\}$$

Sets: Let $\mathbb{E}$ be a collection of sets E

- $\bigcup_{E \in \mathcal{E}} = \{x : x \in E,\ \text{for some } E \in \mathcal{E}\}$
- $\bigcap_{E \in \mathcal{E}} = \{x : x \in E,\ \text{for all } E \in \mathcal{E}\}$
- $\mathcal{E} = \{E_\alpha : \alpha \in A\} = \{E_{\substack{\alpha \\ \alpha \in A}}\}$
- $E_\alpha \cap E_\beta = \emptyset$ for $\alpha \neq \beta \iff E_\alpha$ and E_β are *disjoint*

Limsup and Liminf: For $\{E_n\}_{n=1}^{\infty}$,

$$\limsup E_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n$$
$$\liminf E_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} E_n$$

Exercise: Prove that

$$\limsup E_n = \{x : x \in E_n \text{ for infinitely many } n\}$$
$$\liminf E_n = \{x : x \in E_n \text{ for all but finitely many } n\}$$

i.e. after first finite n, x is in E_n for all n.

Difference and Symmetric Difference: Let E and F be two sets

$$E \setminus F = \{x : x \in E, x \notin F\}$$
$$E \triangle F = (E \setminus F) \cup (F \setminus E)$$
$$E^c = X \setminus E,\ E \subseteq X$$

De Morgan's Laws:

$$\left(\bigcup_{\alpha \in A} E_\alpha \right)^c = \bigcap_{\alpha \in A} E_\alpha^c$$
$$\left(\bigcap_{\alpha \in A} E_\alpha \right)^c = \bigcup_{\alpha \in A} E_\alpha^c$$

Exercise: Prove De Morgan's Laws.

Cartesian Product: If X and Y are sets, then $X \times Y$ is the *ordered* set

$$X \times Y = \{(x, y) : x \in X, y \in Y\}$$

Relations

Relations: A *relation* R from X to Y is a subset of $X \times Y$ such that

$$xRy \iff (x, y) \in R$$

Equivalence relation: A relation $\sim$ is an equivalence relation in the special case $Y = X$ if it is

- Reflexive: $x \sim x \quad \forall x \in X$
- Symmetric $x \sim y \iff y \sim x$
- Transitive $x \sim y, y \sim z \implies x \sim z$

Functions

Mappings: A mapping/function $f : X \to Y$ is a relation R from X to Y such that $\forall x \in X$, there exists a *unique* $y \in Y$ such that xRy. We write $y = f(x)$.

Composition: If $f : X \to Y$ and $g : Y \to Z$, then $g \circ f : X \to Z$ is a function such that $g \circ f(x) = g(f(x))$

Images: If $D \subseteq X, E \subseteq Y$, the *image* of D (and the *inverse image*/pre-image of E) under $f : X \to Y$ is

$$f(D) = \{f(x) : x \in D\}$$
$$f^{-1}(E) = \{x \in X : f(x) \in E\}$$

For $f : X \to Y$ we further call X the *domain* of f and Y the *codomain* of f. The *range/image* of f is $f(X)$.

Inverses: f^{-1} defines an operation on $P(X)$ such that

$$f^{-1}\left(\bigcup_{\alpha \in A} E_\alpha\right) = \bigcup_{\alpha \in A} f^{-1}(E_\alpha)$$
$$f^{-1}\left(\bigcap_{\alpha \in A} E_\alpha\right) = \bigcap_{\alpha \in A} f^{-1}(E_\alpha)$$
$$f^{-1}(E^c) = (f^{-1}(E))^c$$

Exercise: Prove the above properties of inverses. Warning: in general, f also commutes with unions but *not* with intersections. Why?

Bijectivity:

- f is *injective* iff $f(x_1) = f(x_2) \implies x_1 = x_2$
- f is *surjective* iff $\forall y \in Y, \exists x \in X$ s.t. $f(x) = y$
- f is *bijective* iff it is both injective and surjective

In the case of a bijective mapping f, then f^{-1} is a function from Y to X (i.e. f^{-1} has a unique value for bijective f)

Sequences

Sequences: A sequence in a set X is a function $f : \mathbb{N} \to X$. We $\{x_n\}$ for $x_n \in X$

Subsequence: A subsequence $x_{n_k} \subseteq \{x_n\}$ with $n_k \in \{1, \dots, \infty\}$

Ordering

Partial ordering: a partial ordering on a nonempty set X is a relation R on X such that

- If xRy and yRz, then xRz (transitivity)
- If xRy and yRx, then $x = y$ (antisymmetry)
- xRx for all x (reflexivity)

Example: Let E be a set. Consider the relation $\subseteq$. Let $E_1, E_2, E_3 \subseteq E$.

- $E_1 \subseteq E_2$ and $E_2 \subseteq E_3$ implies $E_1 \subseteq E_3$ (transitivity ✓)
- $E_1 \subseteq E_2$ and $E_2 \subseteq E_1$ implies $E_1 = E_2$ (antisymmetry ✓)
- $E_1 \subseteq E_1$ (reflexivity ✓)

Therefore, inclusion (with equality) is a partial ordering. (Proof for first two by considering elements, proof for last by equality)

Total ordering: A total ordering/linear ordering is a partial ordering such that for all $x, y \in X$, either xRy or yRx.

Example: Inclusion is not a total ordering on $P(X)$ since (in general) $E_1 \not\subseteq E_2$ and $E_2 \not\subseteq E_1$ for $E_1 \neq E_2$.

1.2 Sept 10

Recall: a *partial ordering* is a relation that satisfies

1. if xRy and yRz, then xRz
2. if xRy and yRx, then $x = y$
3. xRx for all x

Examples:

- In the real numbers, $\leq$ is the typical ordering.
- For a set X and its power set $P(X)$, $\subseteq$ is a partial ordering.

Warning: In this class, we will use $\leq$ to denote an abstract partial ordering.

Total/Linear Ordering: A total ordering is a partial ordering such that for all $x, y \in X$, either $x \leq y$ or $y \leq x$.

Extrema: If X is partially ordered by $\leq$, a *maximal* (resp. *minimal*) element of X is an element $x \in X$ such that $x \leq y \implies y = x$

Bounds: If $E \subseteq X$, an *upper* (resp. *lower*) *bound* for E is an element $x \in X$ such that $y \leq x$ (resp. $x \leq y$) for all $y \in E$.

Zorn's Lemma (transfinite induction): If X is partially ordered by $\leq$, assume every linearly ordered subset of X has an upper bound. Then X has a maximal element.

Proof: We regard this as axiomatic

Well-Ordering: A set X is *well-ordered* if

1. it is linearly ordered by $\leq$
2. every nonempty subset of X has a minimal element.

Well-ordering Principle: Every non-empty set X can be well-ordered

Proof: Consider $\mathcal{W} = \{\text{all well-ordered subsets of } X\}$.

Suppose there exist well-ordered sets $E_1, E_2 \subseteq W$. Then each has a minimal element.

We know $\mathcal{W}$ is non-empty because for all finite subsets of X, we can order them (using the normal linear order on $\mathbb{R}$).

We will proceed by defining a relation R between the linear orderings $\leq_1$ and $\leq_2$ of E_1 and E_2 respectively. We will say $\leq_1 R \leq_2$ if:

1. $\leq_2$ extends $\leq_1$ (i.e. $E_1 \subseteq E_2$ and $\leq_1 = \leq_2$ on E_1)
2. If $x \notin E_1, x \in E_2$, then $y \leq_2 x$ for all $y \in E_1$

Exercise: Prove that R is a partial ordering in $\mathcal{W}$

Assume $\mathcal{S} = \{\leq_\alpha ; R\}$ is the set of linear orderings $\leq_\alpha$ of $E_\alpha \subseteq \mathcal{W}$ for $\alpha \in A$. Thus, $\leq_\alpha R \leq_\beta$ for $\alpha, \beta \in A$.

Claim: Let

$$E_\infty = \bigcup_{\alpha \in A} E_\alpha$$

equipped with the partial ordering $\leq_\infty$ such that $\leq_\infty \big|_{E_\alpha} = \leq_\alpha$ for all $\alpha \in A$.

Clearly, $\leq_\alpha R \leq_\infty$ for all $\alpha \in A$. Then for any sequence of well-ordered sets in $\mathcal{W}$, E_∞ is an upper-bound.

Exercise: Verify that $\leq_\alpha R \leq_\infty$ is well defined and that E_∞ is an upper bound for $\mathcal{W}$

By Zorn's Lemma, there exists a maximal element $E_{\max} \in \mathcal{W}$. (Verify it's a well-ordering by extending $\leq_{\max}$ to include any $x_0 \in X \setminus E_{\max}$ such that $x \leq x_0$ for all $x \in E_{\max}$).

Consider $E_{\max} \cup \{x_0\}$. Clearly, $E_{\max} \leq E_{\max} \cup \{x_0\}$, so $E_{\max} \cup \{x_0\}$ and by the extension above, $E_{\max} \cup \{x_0\} \in \mathcal{W}$. This contradicts the maximality of $E_{\max}$, so $E_{\max} = X$.

Definition: Let $\prod_{\alpha \in A} X_\alpha$ be the set of all maps $f : A \to \bigcup_{\alpha \in A} X_\alpha$ such that $f(\alpha) \in X_\alpha$ for all $\alpha \in A$.

Axiom of Choice: If $\{X_\alpha\}_{\alpha \in A}$ is a nonempty collection of nonempty sets, $\prod_{\alpha \in A} X_\alpha$ is nonempty, i.e. there exists at least one choice function f

Proof: Let $X = \bigcup_{\alpha \in A} X_\alpha$. Pick a well-ordering on X and $\alpha \in A$. Let $f(\alpha)$ be the minimal element of X_α. Then $f \in \prod_{\alpha \in A} X_\alpha$

Cardinality

Definition:

- card $X \leq$ card Y if there exists an injective function $f : X \to Y$
- card $X =$ card Y if there exists a bijective function $f : X \to Y$
- card $X \geq$ card Y if card $X \leq$ card Y but card $X \neq$ card Y there exists a surjective function $f : X \to Y$

Property: card $X \leq$ card Y iff card $Y \geq$ card X

Proof: card $X \leq$ card Y implies there exists an injective $f : X \to Y$. Pick $x_0 \in X$ and define

$g : Y \to X$ by

$$g(y) = \begin{cases} f^{-1}(y) & y \in f(X) \\ x_0 & \text{otherwise} \end{cases}$$

In the first case, we have injectivity of f so each $f^{-1}(y)$ is unique. In the second case we ensure surjectivity.

Conversely, if $g : Y \to X$ is surjective, consider $g^{-1}(\{x\})$ for $x \in X$. These sets are non-empty and disjoint because f is a map (each x can map to a single y). Then any $f \in \prod_{x \in X} g^{-1}(\{x\})$ is an injection from X to Y.

1.3 Sept 12

Property: For any sets x and Y, either card $X \leq$ card Y or card $Y \leq$ card X

Proof Sketch: Consider the (non-empty) set

$$J = \{\text{all injections } f_E : X \to Y \text{ with respect to } E \subseteq X\}$$

Define a relation R on J such that $f_{E_1} R f_{e_2}$ if $E_1 \subseteq E_2$ and $f_{E_2}\big|_{E_1} = f_{E_1}$, i.e. f_{E_2} is an extension of f_{E_1}.

Repeating the argument of the Well-Ordering Principle, R is a partial ordering.

Then we can find an upper bound for J by considering the union of all $E \in J$ and extending the injections.

By Zorn's Lemma, there exists a maximal element $f_{E_{\max}} \in J$ with respect to the ordering R.

Case 1: Suppose $E_{\max} = X$. Then $f_{E_{\max}}$ is an injection from X to Y so card $X \leq$ card Y

Case 2: Suppose $E_{\max} \subsetneq X$. Then $\exists x_0 \in X \setminus E_{\max}$. Consider the image $f(E_{\max})$. We claim $f(E_{\max}) = Y$ so $f^{-1}_{E_{\max}}$ is defined on all of Y and is injective $Y \to X$ and we are done. Thus, it only remains to show $f(E_{\max}) = Y$.

If the claim is not true, $\exists y_0 \in Y$ but $y_0 \notin f_{E_{\max}}(X)$ but this is a contradiction to maximality (as in the Well-Ordering Principle proof).

Schröder-Bernstein Theorem: If card $X \leq$ card Y and card $Y \leq$ card X, then card $X =$ card Y

Note: This seems trivial but in fact the two functions are not necessarily the same so we must construct our own bijection.

Proof: Denote the cardinality injections $f : X \to Y$ and $g : Y \to X$.

If $f(X) = Y$, then f is a bijection and we are done.

If $f(X) \neq Y$ (i.e. $f(X) \subsetneq Y$), the consider $Y_1 = Y \setminus f(X)$ and $g(Y_1)$. Then $f(Y_1) \subsetneq X$, so call $X_1 = f(Y_1)$. We now have a bijection $X_1 \to Y_1$.

Let's repeat. $f(X \setminus X_1) \subsetneq Y \setminus Y_1$ so define $Y_2 = (Y \setminus Y_1) \setminus f(X \setminus X_1)$.

Now we know $f(X_1) \subseteq Y_2$ and $f^{-1}(Y_1) \subseteq X_1$ so we can define a bijection $X_2 \to Y_2$.

Assume $X_1, \dots X_n$ and $Y_1, \dots, Y_n$ are constructed. WLOG assume that this procedure can be repeated infinitely (or else we would already have a bijection).

Define
$$\left(Y \setminus \bigcup_{i=1}^{n} Y_i\right) \setminus f\left(X \setminus \bigcup_{i=1}^{n} X_i\right) = Y_{n+1}$$
since $f(X_i) \subseteq Y_{i+1}$.

Exercise: Verify that
$$g : \bigcup_{i=1}^{\infty} Y_i \to \bigcup_{i=1}^{\infty} X_i$$
is a bijection and further that
$$f : \left(X \setminus \bigcup_{i=1}^{\infty} X_i\right) \to \left(Y \setminus \bigcup_{i=1}^{\infty} Y_n\right)$$
is also a bijection.

Together, these steps show that we have a bijection on the full sets X and Y.

Proposition: For any set X, card $X <$ card $P(X)$

Proof: Clearly, $\forall x \in X$, we have an injection $f : X \hookrightarrow P(X)$ defined by $f(x) = \{x\}$.

We claim there is no surjection $g : X \to P(X)$ and proceed by contradiction.

Let $g : X \twoheadrightarrow P(X)$. Define
$$Y = \{x \in X \text{ s.t. } x \notin g(x)\}$$

We claim $Y \notin g(X)$. If not, assume $x_0 \in X$ such that $g(x_0) = Y$.

Case 1: If $x_0 \in Y$, then $x_0 \notin g(x_0) = Y$ - contradiction

Case 2: If $x_0 \notin Y$, then $x_0 \in g(x_0) = Y$ - contradiction

Therefore, $Y \notin g(X)$ so g is not surjective.

Countable: A set X is *countably infinite* if card $X \leq$ card $\mathbb{N}$.

Proposition:

(a) If X and Y are countable, so is $X \times Y$.

(b) If A is countable and X_α is countable for every $\alpha \in A$, then $\bigcup_{\alpha \in A} X_\alpha$ is countable.

Proof:

(a) card $X =$ card $Y =$ card $\mathbb{N}$ so it suffices to show $\mathbb{N} \times \mathbb{N} =$ card $\mathbb{N}$

$\forall n \in \mathbb{N}$, define $f(n) \hookrightarrow (n, 1) \in \mathbb{N} \times \mathbb{N}$.

Consider $g((m,n)) \to 2^m 3^n \in \mathbb{N}$. Is this injective? Consider $g(m_1,n_1) = 2^{m_1}3^{n_1}$. By the unique prime factorization of integers, $2^{m_1}3^{n_1} = 2^m 3^n$ iff $(m_1,n_1) = (m,n)$ so g is injective.

Now we can use Schröder-Bernstein and we are done.

(b) As A is countable, $\forall \alpha \in A$, $\exists f_\alpha : \mathbb{N} \to X_\alpha$ So we can define $F : \mathbb{N} \times A \to \bigcup_{\alpha \in A} X_\alpha$ by

$$F(n,\alpha) = f_\alpha(n)$$

which is surjective

Corollary: $\mathbb{Z}$ and $\mathbb{Q}$ are countable

Proof: $\mathbb{Z} = \mathbb{N} \cup \{-\mathbb{N}\} \cup 0$

We can define $f : \mathbb{Z}^2 \to \mathbb{Q}$ by

$$f(m,n) = \begin{cases} \frac{m}{n} & n \neq 0 \\ 0 & n = 0 \end{cases}$$

Convention for this course: We will use $\mathbb{R}$ to denote the standard reals and will define the *extended reals* $\overline{R}$ by $\mathbb{R} \cup \pm\infty$

Under this notation, we can state that for any $E \subseteq \overline{R}$, $\sup \overline{E}$ and $\inf \overline{E}$ are always well-defined, i.e. all sets are bounded above by ∞ and below by $-\infty$.

We define the following rules:

- $X \pm \infty = \pm\infty$
- $\infty + \infty = \infty$
- $-\infty - \infty = -\infty$
- $\infty - \infty$ is undefined
- $x(\pm\infty) = \pm\infty$ for $x > 0$ and $x(\pm\infty) = \mp\infty$ for $x < 0$
- $0 \cdot (\pm\infty) = 0$

Note: this last point does *not* talk about limits, it is just notation

Proposition: Every open set in $\mathbb{R}$ is a countable disjoint union of open intervals

Proof Sketch: For all $x \in U$, there exists an open interval $I_{\alpha,\beta} = (\alpha,\beta) \subseteq U$ with $\alpha < x < \beta$.

Let $\mathcal{J}_x = \{x \in I_{\alpha,\beta} \mid I_{\alpha,\beta} \in U\}$.

Take $\alpha_{\inf} = \inf \alpha$ and $\beta_{\sup} = \sup \beta$.

Exercise: Check that $x \in (\alpha_{\inf}, \beta_{\sup}) \subseteq U$

We call $I_x = (\alpha_{\inf}, \beta_{\sup})$ for all $x \in U$

We claim $\forall x, y \in U$, either $I_x \cap I_y = \emptyset$ or $I_x = I_y$.

Suppose $I_x \cap I_y \neq \emptyset$. Then $I_x \cup I_y$ is an open interval containing x, so $I_x \cup I_y \in \mathcal{J}_x$ but I_x is maximal so this is a contradiction unless $I_x = I_y$

Now we can write
$$U = \bigcup_{x \in U} I_x$$

Why is this countable? We can define an injection $U \to \mathbb{Q}$ by choosing a rational number in each I_x (exist by density of $\mathbb{Q}$).

Metric Spaces

Definition: A *metric space* is a set X together with a *distance function* $\rho : X \times X \to [0, \infty)$ such that

1. $\rho(x, y) = 0 \iff x = y$
2. $\rho(x, y) = \rho(y, x)$
3. $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$

Examples:

- $\mathbb{R}^n$ with $\rho(x, y) = |x - y|$
- Set of continuous functions f over $[0, 1]$ with $\rho_1(f, g) = \int 0^1 |f(x) - g(x)| \, dx$ (or alternatively $\rho_\infty = \sup_{0 \leq x \leq 1} |f(x) - g(x)|$)

Exercise: Check the above are metric spaces

1.4 Sept 17

Closed and Open Sets

Open ball: Let (X, ρ) be a metric space. If $x \in X$, $r > 0$, we define the *open ball* $B(x, r) = \{y \in X \text{ s.t. } \rho(x, y) < r\}$

Open set: a set E is open iff $\forall x \in E, \exists r > 0$ s.t. $B(x, r) \subseteq E$

Closed set: a set E is closed iff E^c is open

Example: $B(x, r)$ is open. Consider $y \in B(x, r)$. Then $\rho(x, y) = s < r$. By the triangle inequality, $B(y, r - s) \subseteq B(x, r)$

Exercise: Prove that $B(x, r)$ is open

Properties:

- $\emptyset$ is open
- If U_x are open sets, $\bigcup_{x \in A} U_x$ is open (as is the finite intersection)
- If F_x are closed sets, $\bigcap_{x \in A} F_x$ is closed (as is the finite union)

Interior: Let $E \subseteq X$. The *interior* of E is
$$\mathring{E} = \bigcup_{O \subseteq E} O$$

(this is the largest open set in E)

Closure: The *closure* of E is

$$\overline{E} = \bigcap_{E \subseteq F} F$$

(this is the smallest closed set containing E)

Proposition: Let (X, ρ) be a metric space. Let $E \subseteq X$ and $x \in X$. Then the following are equivalent:

(a) $x \in \overline{E}$

(b) $B(x, r) \cap E \neq \emptyset$ for all $r > 0$

(c) $\exists (x_n) \subseteq E$ such that $x_n \to x$

Proof: $((a) \to (b))$ Let $x \in \overline{E}$. Suppose $\exists r_0 > 0$ such that $B(x, r) \cap E = \emptyset$. Then $E \subseteq (B(x, r_0))^c$. But $(B(x, r_0))^c$ is closed so $\overline{E} \subseteq (B(x, r_0))^c$ so $x \in B(x, r_0) \subseteq (\overline{E})^c$ but this implies $x \in (\overline{E})^c$ which is a contradiction.

$((b) \to (c))$ Let $r = \frac{1}{n}$. By (b), $B(x, \frac{1}{n}) \cap E \neq \emptyset$. Choose $x_n \in B(x, \frac{1}{n}) \cap E$. Certainly $\rho(x_n, x) < \frac{1}{n}$ so $\lim \rho(x_n, x) = 0$ and $x_n \to x$

$((c) \to (a))$ If $x \notin \overline{E}$, $x \in (\overline{E})^c$ but $(\overline{E})^c$ is open so $\exists r > 0$ s.t. $B(x, r) \subseteq (\overline{E})^c \subseteq E^c$. Then there cannot exist any sequence in E. But this contradicts $x_n \to x$

Density

Dense: E is dense in X if $\overline{E} = X$ (examples $\mathbb{R}^n, \mathbb{Q}^n$)

Nowhere dense: E is nowhere dense if $(\overline{E})^\circ = \emptyset$ (example: emptyset)

Separable: X is separable if there exists a countable dense subset $E \subseteq X$

Limits: In this class, $x_n \to x$ iff $\lim_{n\to\infty} \rho(x_n, x) = 0$

Continuity

Let $\mathcal{C} = \{\text{continuous functions on } [0,1]\}$.

Continuity at a point: If (X_1, ρ_1) and (X_2, ρ_2) are metric spaces, $f : X_1 \to X_2$ is continuous at $x \in X_1$ if $\forall \varepsilon > 0, \exists \delta_x > 0$ such that $\forall y \in X_1$ such that $\rho_1(x, y) < \delta_x$ (i.e. $y \in B_1(x, \delta_x)$),

$$\rho_2(f(x), f(y)) < \varepsilon$$

(i.e. $f(y) \in B_2(f(x), \varepsilon)$)

Continuity on a set: f is continuous in X iff f is continuous at every $x \in X$

Uniform Continuity: f is uniformly continuous if δ is independent of x, i.e. $\forall \varepsilon > 0$, $\exists \delta > 0$ such that

$$\rho_1(x, y) < \delta \implies \rho_2(f(x), f(y)) < \varepsilon$$

for all $x \in X$.

Proposition: $f : X_1 \to X_2$ is continuous iff $f^{-1}(U) \subseteq X_1$ is open for all open $U \subseteq X_2$

Proof: Let f be continuous and $U \subseteq X_2$ be open. $f^{-1}(U) = \emptyset$ is open so take $x \in f^{-1}(U)$. Then $f(x) = y \in U$.

Since U is open, $\exists \varepsilon_y > 0$ s.t. $B_2(y, \varepsilon_y) = B_2(f(x), \varepsilon_y) \subseteq U$.

By continuity, $\exists \delta_x > 0$ such that $\forall z \in B_1(x, \delta_2)$,

$$\rho_2(f(x), f(z)) < \varepsilon_y \implies f(z) \in B_2(y, \varepsilon_y) \subseteq U \implies z \in f^{-1}(U)$$

so $B_1(x_1, \delta_x) \subseteq f^{-1}(U)$ and $f^{-1}(U)$ is open.

Conversely, suppose $f^{-1}(U)$ is open for all open $U \subseteq X_2$. Let $\varepsilon > 0$. Consider $y = f(x) X_2$. Then $B_2(y, \varepsilon)$ is open so $f^{-1}(B_2(y, \varepsilon))$ is open by assumption.

Let $x \in f^{-1}(B_2(y, \varepsilon))$. Then $\exists \delta_x$ such that $B_1(x, \delta_x) \subseteq f^{-1}(B_2(y, \varepsilon))$.

Then $f(B_1(x, \delta_x)) \subseteq B_2(y, \varepsilon)$ which is precisely the definition of continuity.

Cauchy Sequences

Cauchy Sequence: A sequence (x_n) in a metric space (X, ρ) is Cauchy if $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ such that $\forall m, n \geq N$,

$$\rho(x_m, x_n) < \varepsilon$$

Completeness: A subset $E \subseteq X$ is *complete* if every Cauchy sequence $x_n \in E$ has a limit $x \in E$

Examples:

- In $\mathbb{R}^n$, any bounded closed subset is complete.
- $(\mathcal{C}, \rho_\infty)$ is complete

Exercise: Prove that $(\mathcal{C}, \rho_\infty)$ is complete for

$$\rho_\infty(x, y) = \sup_{x \in [0,1]} |f(x) - g(x)|$$

(though in general this is not true for other metrics)

Proposition: A closed subset (X, ρ) of a complete metric space is complete and complete subsets of a metric space must be closed

Proof:

Exercise

Set Distance:

- Let $x \in X$ and $E \subseteq X$. The *distance* from x to E is

$$\rho(x, E) = \inf\{\rho(x, y) : y \in E\}$$

- For $E, F \subseteq X$,
$$\rho(E, F) = \inf\{\rho(x, y) : x \in E, y \in F\}$$

Diameter: diam $E = \sup\{\rho(x, y) : x, y \in E\}$

Bounded: E is bounded iff diam $E < \infty$

Open cover: Let $\{V_\alpha\}_{\alpha \in A}$ be a family of sets. $\{V_\alpha\}$ *covers* E if

$$E \subseteq \bigcup_{\alpha \in A} V_\alpha$$

Total boundedness: E is *totally bounded* if $\forall \varepsilon > 0$, E can be covered by finitely many balls of radius ε

Example: $\mathbb{R}^n$ is totally bounded. *Proof:* consider a hypercube of side length R. Clearly we can divide this into ε-cubes and then take slightly larger balls to cover the whole space.

Theorem (Characterization of Compactness): The following are equivalent:

1. E is complete and totally bounded
2. Every sequence in E has a convergent subsequence with its limit in E
3. If $\{V_\alpha\}_{\alpha \in A}$ is an open cover of E, then there exists a finite set $F \subseteq A$ such that $\{U_\alpha\}_{\alpha \in F}$ covers E

Proof: HW

1.5 Sept 19

Products of Metric Spaces: Let (X, ρ_1) and (Y, ρ_2) be metric spaces. Define the *product metric* on $X \times Y$ by $(X_1 \times X_2, \rho_1 \times \rho_2)$ where

$$\rho_1 \times \rho_2 = \sqrt{\rho_1^2(x_1, y_1) + \rho^2(x_2, y_2)}$$

(so called *Euclidean Metric*)

Though many other metrics are possible, such as $\max(\rho_1, \rho_2)$ and $\rho_1 + \rho_2$.

In general, we will simply take the Euclidean metric because all these metrics are equivalent in the sense that $\exists C_1, C_2$ such that
$$C_1(\rho_1 \times \rho_2)_1 \leq C_2(\rho_1 \times \rho_2)_2 \leq C_2(\rho_1 \times \rho_2)_3$$

Properties:

- $\rho_1 \times \rho_2 \to 0 \iff \rho_1 \to 0$ and $\rho_2 \to 0$

Chapter 2

Measure Theory

2.1 Sept 19

Measure Theory Motivation

Riemann Integral: Let $f : [a, b] \to \mathbb{R}$. We subdivide $[a, b]$ by

$$a = x_0 < x_1 < \cdots < x_n = b$$

and define subintervals $[x_i, x_{i+1}]$.

Then

$$\int f(x)\, dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i) \cdot (x_{i+1} - x_i)$$

Convergence: Many times, we are interested in the question:

$$\lim_{n\to\infty} \int_0^1 f_n(x)\, dx \stackrel{?}{=} \int_0^1 f(x)\, dx$$

for $f_n(x) \to f(x)$.

This is easy when $f_n \to f$ uniformly but in general, we need something else.

In Riemann integration, we divide the domain into intervals and sum the function over these intervals.

In Lebesgue integration, we instead divide *the range*, i.e. we take a set

$$E_i = \{x : a_n \le f(x) \le a_{n+1}\}$$

Measure: We define $\mu(E)$, the *measure* of a subset, by:

1. (Countable Additivity) $\{E_n\}$ such that $E_i \cap E_j = \emptyset$ for $i \ne j$ then $\mu(\bigcup_{n=1}^{\infty} E_n) = \sum_{i=1}^{n} \mu(E_n)$
2. (Translation invariance) $\mu(E + r) = \mu(\{x + r : x \in E\}) = \mu(E)$
3. $\mu([0, 1]) = 1$

Proposition: There is no measure μ satisfying the above properties which is defined for all subsets of $[0, 1)$

Proof: Step 1. Let $\mathbb{Q}_1 = \mathbb{Q} \cap [0, 1)$. Define an equivalence relation $x \sim y$ iff $x - y \in \mathbb{Q}_1$.

Now consider the equivalence class $\mathcal{E}_x = \{y \in [0, 1) : y \sim x\}$. (As it is an equivalence class: $\mathcal{E}_x \cap \mathcal{E}_y \neq \emptyset \implies \mathcal{E}_x = \mathcal{E}_y$)

And clearly,

$$[0, 1) = \bigcup_{x \in [0,1)} \mathcal{E}_x$$

By the Axiom of Choice, choose a unique element $e_x \in \mathcal{E}_x$. Define $N = \{e_x\}$. Now $e_x - e_y \notin \mathbb{Q}_1$.

Step 2. $\forall r \in \mathbb{Q}_1$, define

$$N_r = \{e_x + r : e_x \in N \cap [0, 1 - r)\} \cup \{e_x + r - 1, e_x \in N \cap [1 - r, 1]\}$$

(the first set is the points that don't leave the interval under translation, the second set is the pullback of the points that do)

Step 3. We claim

$$[0, 1) = \bigcup N_r,\ N_r \cap N_s = \emptyset \text{ for } r \neq s$$

Proof:

1. (Subset) $\forall y \in [0, 1)$, $\exists e_x \in N$ such that $y - e_x \in \mathbb{Q}_1$.

   If $y \geq e_x$, $r = e_x - y + 1$. Otherwise, $r = e_x - y$.

2. (Disjoint Union) Suppose $N_r \cap N_s \neq \emptyset$. Let $r \neq s$. Select $y \in N_r \cap N_s$ so $y - s \in N$ and $y - r \in N$

   Case 1. $y - s \neq y - r$. But then

   $$(y - r) - (y - s) = s - r \in \mathbb{Q}_1$$

   which is a contradiction of the construction of N.

   Case 2. $y - s \neq y - r + 1$. Contradiction again by rational difference.

Step 4. By the definition of a measure,

$$\begin{aligned} \mu(N_r) &= \mu(N_r \cap (0, 1 - r)) + \mu(N_r \cap [1 - r, 1]) \\ &= \mu(N) \end{aligned}$$

Exercise: Check that $\mu(N_r) = \mu(N)$

But by countable Additivity,

$$1 = \mu([0, 1)) = \sum_{r \in \mathbb{Q}_1}^{\infty} \mu(N_r) = \begin{cases} 0 \\ \infty \end{cases}$$

which is a contradiction.

Conclusion: it is not always possible to define a measure so we need to be careful.

Algebras

Algebra: Given a set X, an *algebra* is a collection of subsets $\mathcal{A} \subseteq P(X)$ such that if $E_1, \dots, E_n \subseteq \mathcal{A}$,

1. $\bigcup_{i=1}^{n} E_i \in \mathcal{A}$
2. $E \in \mathcal{A} \implies E^c \in \mathcal{A}$

Property 2 gives us that $X \in \mathcal{A}$ and $\emptyset \in \mathcal{A}$ ($E \cup E^c = X,\ X^c = \emptyset$)

Sigma Algebra: An algebra $\mathcal{A}$ is a *σ-algebra* if it is closed under countable unions and complements, i.e. for $E_1, E_2, \dots \in \mathcal{A}$,

1. $\bigcup_{i=1}^{\infty} E_i \in \mathcal{A}$
2. $E \in \mathcal{A} \implies E^c \in \mathcal{A}$

Remark: It suffices to demand closure for disjoint countable unions since

$$\bigcup_{n=1}^{\infty} E_i = \bigcup_{n=1}^{\infty} F_i$$

for $F_k = E_k \setminus \bigcup_{i=1}^{k-1} E_i$ and $F_i \cap F_{i+1} = \emptyset$

Examples:

- $P(X)$
- ϕ, X
- $\mathcal{A} = \{E \subseteq X : E \text{ countable or } E^c \text{ countable}\}$

Proposition: Let $\mathcal{A}_1, \mathcal{A}_2$ be two σ-algebras on X. Then $\mathcal{A}_1 \cap \mathcal{A}_2$ is also a σ-algebra

Exercise: Prove this proposition (easy using definition)

Generated σ-algebra: Given a collection of subsets $\mathcal{E} \subseteq P(X)$, there exists a smallest σ-algebra containing $\mathcal{E}$, denoted

$$M(\mathcal{E}) = \bigcap_{\mathcal{E} \subseteq \mathcal{A}} \mathcal{A}$$

Lemma: $\mathcal{E} \subseteq M(\mathcal{F}) \implies M(\mathcal{E}) \subseteq M(\mathcal{F})$

Proof: Omitted

Metric Spaces

Borel σ-algebra: Let (X, ρ) be a metric space. We call the σ-algebra generated by the open sets of X, the *Borel σ-algebra* B_x on X.

This is a σ-algebra because $X, \emptyset, \bigcup_{i=A} U_i$ are all open since their union is open and we have complements from the generating set.

We define

$$\bigcup_{n=1}^{\infty} F_n = F_\sigma$$
$$\bigcap_{n=1}^{\infty} O_n = G_\delta$$

for F_n closed and O_n open.

Example: The Borel set of $\mathbb{R}$, $B_\mathbb{R}$ can be generated by any of the following:

1. open intervals $\mathcal{E}_1 = \{(a, b) : a < b\}$
2. closed intervals $\mathcal{E}_2 = \{[a, b] : a < b\}$
3. the half-open intervals $\mathcal{E}_3 = \{(a, b] : a < b\}$, $\mathcal{E}_4 = \{[a, b) : a < b\}$
4. open rays $\mathcal{E}_5 = \{(a, \infty) : a \in \mathbb{R}\}$, $\mathcal{E}_6 = \{(-\infty, a) : a \in \mathbb{R}\}$
5. closed rays

Exercise:

1. Prove that $(a, b] = \bigcap_{n=1}^{\infty}(a, b + \frac{1}{n})$ and $(a, b) = \bigcup_{n=1}^{\infty}[a + \frac{1}{n}, b - \frac{1}{n}]$
2. Prove that the above methods all generate $B_\mathbb{R}$

Conclusion: any open set in $\mathbb{R}$ is the countable union of open intervals

2.2 Sept 24

Recall last time, we were trying to characterization the Borel σ-algebra of $\mathbb{R}$, $\mathcal{B}_\mathbb{R}$.

Proposition: We claim that $\mathcal{B}_\mathbb{R}$ is generated by:

1. open intervals $\mathcal{E}_1 = \{(a, b) : a < b\}$
2. closed intervals $\mathcal{E}_2 = \{[a, b] : a < b\}$
3. the half-open intervals $\mathcal{E}_3 = \{(a, b] : a < b\}$, $\mathcal{E}_4 = \{[a, b) : a < b\}$
4. open rays $\mathcal{E}_5 = \{(a, \infty) : a \in \mathbb{R}\}$, $\mathcal{E}_6 = \{(-\infty, a) : a \in \mathbb{R}\}$
5. closed rays

Proof:

1. Open intervals.

   Let $\mathbb{E}_1 = \{(a, b) : a < b\}$. Clearly $B_{E_1} \subseteq B_\mathbb{R}$ because any open set $O \subseteq \mathcal{B}_\mathbb{R}$.

   For the other direction, we also have

   $$O = \bigcup_{i=1}^{\infty}(a_i, b_i)$$

   (a countable union), so $B_\mathbb{R} \subseteq B_{\mathcal{E}_1}$
2. Closed intervals.

We claim

$$(a,b) = \bigcup_{n=1}^{N} [a + \frac{1}{n}, b - \frac{1}{n}]$$

for N sufficiently large.

Proof: HW

Now $\forall y \in (a,b)$,

$$y \in \bigcup_{n=1}^{N} [a + \frac{1}{n}, \frac{b}{\frac{1}{n}}] \implies a < y < b$$

for N sufficiently large.

For the other direction, take $[a,b] \in \mathcal{E}_2$. Then

$$[a,b] = \bigcap_{n=1}^{\infty} (a - \frac{1}{n}, b + \frac{1}{n})$$

Proof: $[a,b] \subseteq \bigcap_{n=1}^{\infty}(a - \frac{1}{n}, b + \frac{1}{n})$ is clear.

For the other direction, let $yin \bigcap_{n=1}^{\infty}(a - \frac{1}{n}, b + \frac{1}{n})$. Suppose $a \leq y \leq b$ is false. Then $y \notin (a - \frac{1}{N}, b + \frac{1}{N})$ so it cannot be in the intersection

Exercise: Prove the last two versions: half intervals and rays

Recall: For a cartesian product $X_1 \times X_2 \times \cdots \times X_n$ of metric spaces with (X_i, ρ_i), we define the *product metric* by $(X_1 \times X_2 \times \cdots \times X_n, \rho)$ where

$$\rho(\overline{x}, \overline{y}) = \sqrt{\rho_1^2(x_1, y_1) + \cdots + \rho_n^2(x_n, y_n)}$$

where $\overline{x} = (x_1, x_2, \ldots, x_n)$ with $x_i \in X_i$ (and similarly for $\overline{y}$)

Proposition:

$$\lim_{m\to\infty} \rho(\overline{x}, \overline{y}) = 0 \iff \lim_{m\to\infty} \rho_i(x_i^m, y_i^m) = 0$$

Proof: Omitted

In this way, we can consider $\mathbb{R}^n$ as a metric space with this Euclidean metric. What is the Borel set of $\mathbb{R}^n$?

Proposition: $B_{\mathbb{R}^n}$ is

Proof: First take O_i open set in X_i

$$\bigoplus_{i=1}^{n} O_i = O_1 \times O_2 \times \cdots \times O_n$$

We claim that this is an open set in the $X_1 \times X_2 \times \cdots \times X_n$ topology.

Proof: Take $\overline{x} \in \bigoplus_{i=1}^n O_i$ with $x_i \in O_i$.

It suffices to show $\exists \varepsilon_0 > 0$ such that $B_{\varepsilon_0}(\overline{x}) \subseteq \bigoplus_{i=1}^n O_i$ where

$$B_{\varepsilon_0}(\overline{x}) = \{\overline{y} : \rho(\overline{x}, \overline{y}) < \varepsilon_0\}$$

so $\overline{y} \in B_\varepsilon(\overline{x})$ iff $\rho_i(x_i, y_i) < \varepsilon_0$ for all i.

Hence $y_i \in B_{\varepsilon_0}(x_i) \subseteq O_i$

Let $\bigotimes_{i=1}^n \mathcal{B}_{x_i}$ be the Borel set generated by $\bigoplus_{i=1}^n O_i$

Clearly, $\bigoplus_{i=1}^n \mathcal{B}_{x_i} \subseteq \mathcal{B}_{x_1 \times x_2 \times \cdots \times x_n}$

Lemma: If x_i is separable then

$$\bigotimes_{i=1}^n B_{x_i} = \mathcal{B}_{x_1 \times x_2 \times \cdots \times x_n}$$

In particular:

$$\bigotimes_{i=1}^n \mathcal{B}_{\mathbb{R}} = \mathcal{B}_{\mathbb{R}^n}$$

Proof: It suffices to show that $\forall \overline{x}, \varepsilon$,

$$\mathcal{B}_\varepsilon(\overline{x}, \varepsilon) \subseteq \bigotimes_{i=1}^n B_{x_i}$$

Let $\mathcal{C}_i$ be a countable subset of X_i such that $\overline{\mathcal{C}_i} = X_i$ for all $1 \leq i \leq n$

We claim

$$B_\varepsilon(\overline{x}) \subseteq \bigcup_{\substack{c_i \in \mathcal{C}_i \\ r_i \in \mathbb{Q}}} \bigotimes_{i=1}^n B_{r_i}(c_i)$$

for $\sqrt{r_1^2 + \cdots + r_n^2} < \varepsilon$

(And this has cardinality $\mathbb{N}^{2n}$ so countable)

Pick

$$\overline{y} \in B_\varepsilon(\overline{x}) \subseteq \bigcup_{\substack{c_i \in \mathcal{C}_i \\ r_i \in \mathbb{Q}}} \bigotimes_{i=1}^n B_{r_i}(c_i) \subseteq \bigotimes_{i=1}^n \mathcal{B}_{x_i}$$

Then

$$\sigma(\overline{x}, \overline{y}) = \sqrt{\rho_1^2(y_1, x_2), \ldots, \rho_n^2(y_n, x_n)} < \varepsilon$$

but each $\rho_i^2(y_i, x_i)$ is fixed so we may choose $c_i \in \mathcal{C}, r_i \in \mathbb{Q}$ such that

$$\rho_i(y_i, c_i) < r_i = \rho_i(y_i, x_i) - [\rho(y_i, x_i) - \rho(y_i, c_i)]$$

by density (from separability)

Since $\mathbb{Q}^n \subseteq \mathbb{R}^n$ which is countable and dense, $\mathbb{R}^n$ is separable and we are done.

Measure Spaces

Recall that we could not always define a measure except on a σ-algebra. Therefore, we limit our attention.

Measure space: $(X, \mathcal{M})$ where X is a set and $\mathcal{M}$, a σ-algebra, is the "measurable sets"

Measure: For a measure space $(X, \mathcal{M})$, we define a *measure* $\mu : \mathcal{M} \to [0, \infty]$ such that

1. $\mu(\emptyset) = 0$
2. (Countable additivity) if $\{E_j\}_1^\infty$ is a sequence of pairwise disjoint sets in $\mathcal{M}$, then

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu(E_j)$$

Intuitively, this countable additivity property lets us pull out the limits:

$$\mu\left(\lim_{n\to\infty} \bigcup_{1}^{n} E_j\right) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(E_j)$$

σ-finite: If $\mu(X) = \infty$ but

$$X = \bigcup_{i=1}^{\infty} X_i$$

where $\mu(X_i) < \infty$ for all i, then we call X *σ-finite*

Example: Let $(X, P(X))$ be a measure space. Let $f : X \to [0, \infty]$. For each $E \in P(X)$, we define

$$\mu(E) = \sum_{x\in E} f(x) = \sup\{\sum_{x\in F} f(x) : F \subseteq E \wedge F \text{ finite}\}$$

Exercise: Prove that μ is a measure on $P(X)$

In particular:

- $f(x) = 1$ for all x, then $\mu(E)$ is the *counting measure*
- Take $x_0 \in X$ and define

$$f(x) = \begin{cases} 1 & x = x_0 \\ 0 & x \neq x_0 \end{cases}$$

  is the *Dirac-Delta Mass* at x_0

Example: Let X be an uncountable set. Let $\mathcal{M} = \{E \text{ is finite or } E^c \text{ is finite}\}$

Define

$$\mu(E) = \begin{cases} 0 & E \text{ is countable} \\ 1 & E^c \text{ is countable} \end{cases}$$

Exercise: Check that $\mathcal{M}$ is a σ-algebra and that μ is a measure

2.3 Sept 26

Theorem (Properties of Measures): Let $(X, \mathcal{M}, \mu)$ be a measure space. Then:

1. (Monotonicity) with $E \subseteq F$ with $E, F \in \mathcal{M}$, then
$$\mu(E) \leq \mu(F)$$
2. (Subadditivity) If $\{E_j\}_1^\infty \in \mathcal{M}$, then
$$\mu\left(\bigcup_{j=1}^\infty E_j\right) \leq \sum_{j=1}^\infty \mu(E_j)$$
3. (Continuity from Below) If $\{E_j\}_1^\infty \subseteq \mathcal{M}$ and $E_1 \subseteq E_2 \subseteq \dots$, then
$$\mu\left(\bigcup_{j=1}^\infty E_j\right) = \lim_{j\to\infty} \mu(E_j)$$
4. (Continuity from Above) If $\{E_j\}_1^\infty \subseteq \mathcal{M}$ and $E_1 \supset E_2 \supset \dots$ and $\mu(E_1) < \infty$, then
$$\mu\left(\bigcap_{j=1}^\infty E_j\right) = \lim_{j\to\infty} \mu(E_j)$$

Proof: (1) Take $E \subseteq F \in \mathcal{M}$. We want to use finite additivity. Consider $F \setminus E = F \cap E^c$. Certainly, $E \cap (F \cap E^c) = \emptyset$ so
$$\mu(F) = \mu(E \cup F \setminus E) = \mu(E) + \mu(F \setminus E)$$
but the measure is nonnegative so $\mu(E) \leq \mu(F)$.

And in fact, if $\mu(F) < \infty$, then $\mu(F) - \mu(E) = \mu(F \setminus E)$.

---

(2) Once again, we would like to take advantage of finite additivity by expressing $\bigcup_{j=1}^\infty E_j$ as a countable disjoint union.

Let
$$F_k = E_k \setminus \bigcup_{i=1}^{k-1} E_i = \bigcup_{j=1}^\infty F_k$$
so
$$\mu\left(\bigcup_{j=1}^\infty E_j\right) = \mu\left(\bigcup_{j=1}^\infty F_j\right) = \sum_{j=1}^\infty \mu(F_j) \leq \sum_{j=1}^\infty \mu(E_j)$$

---

(3) Let $E_1 \subseteq E_2 \subseteq \ldots \subseteq E_n \subseteq \ldots$. Denote $E_0 = \emptyset$. Then

$$\begin{aligned} E_1 &= E_1 \setminus \emptyset \\ E_2 &= E_1 \cup (E_2 \setminus E_1) \\ E_3 &= E_2 \cup (E_3 \setminus E_2) \\ &= E_1 \cup (E_2 \setminus E_1) \cup (E_3 \setminus E_2) \\ &= (E_1 \setminus \emptyset) \cup (E_2 \setminus E_1) \cup (E_3 \setminus E_2) \end{aligned}$$

and each of these sets are disjoint.

Inductively define

$$E_n = \bigcup_{k=0}^{n-1} E_{k+1} \setminus E_k$$

We claim

$$\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=0}^{\infty} (E_n \setminus E_{n-1})$$

By additivity,

$$\begin{aligned} \mu\left(\bigcup_{i=1}^{\infty} E_n\right) &= \sum_{n=0}^{\infty} \mu(E_{n+1} \setminus E_n) & &= \lim_{N\to\infty} \sum_{n=0}^{N} \mu(E_{n+1} \setminus E_n) \\ &= \lim_{n\to\infty} \mu(E_n) \end{aligned}$$

---

(4) Let $E_1 \supset E_2 \supset \ldots$ and $\mu(E_1) < \infty$. Define $F_j = E_1 \setminus E_j$.

Clearly, $F_n \subseteq F_{n+1}$. By part 3,

$$\mu\left(\bigcup_{n=1}^{\infty} F_n\right) = \lim_{n\to\infty} \mu(F_n)$$

and

$$F_j = E_1 \setminus E_j \implies \bigcup_{n=1}^{n} F_j = E_1 \setminus \bigcup_{n=1}^{n} E_j$$

(by $E_1 \supset E_2 \supset \ldots$)

So

$$\mu\left(E_1 \setminus \bigcap_{j=1}^{\infty}\right) = \lim_{n\to\infty} \mu(E_1 \setminus \bigcap_{j=1}^{n} E_j)$$

By Part 1,

$$\begin{aligned}
\lim_{n\to\infty} \mu(E_1 \setminus \bigcap_{j=1}^{n} E_j) &= \lim_{n\to\infty} \left[\mu(E_1) - \mu(\bigcap_{j=1}^{n} E_j)\right] \\
&= \lim_{n\to\infty} [\mu(E_1) - \mu(E_n)] \\
&= \lim_{n\to\infty} \mu(E_n)
\end{aligned}$$

Constructing Measures

We have shown that finding measures is hard in general. Let's construct them instead.

Outer Measure: Let X be a set and $\mu^* : P(X) \to [0, \infty]$ be an outer measure if

1. $\mu^*(\emptyset) = 0$
2. (Monotonicity) $E \subseteq F \implies \mu^*(E) \leq \mu^*(F)$
3. (Subadditivity) $\mu^* \left(\bigcup_{j=1}^{\infty} E_j\right) \leq \sum_{j=1}^{\infty} \mu^*(E_j)$

(note: this is *almost* a measure and would be if we allowed additivity rather than subadditivity)

Proposition: Let $\mathcal{E} \subseteq P(X)$ such that $X, \emptyset \in \mathcal{E}$. Define $\rho : \mathcal{E} \to [0, \infty]$ with $\rho(\emptyset) = 0$. $\forall A \subseteq P(X)$, let

$$\mu^*(A) = \inf \left\{ \sum_{i=0}^{\infty} \rho(E_j) \middle| E_j \in \mathcal{E},\ A \subseteq \bigcup_{j=1}^{\infty} E_j \right\}$$

(i.e. take the inf of the sum of all coverings of A). Then μ^* is an outer measure.

Proof:

First note that μ^* is well-defined: certainly $A \subseteq X$ so the set will not be empty and the inf is well defined.

Clearly, $\mu^*(\emptyset) = 0$ because $\rho(\emptyset) = 0$.

(Monotonicity) Let $A \subseteq B$ and $\{E_j \in \mathcal{E}\}_1^{\infty}$ be any covering of B. Since $A \subseteq B$,

$$\mu^*(A) \leq \sum_{n=0}^{\infty} \rho(E_j)$$

Taking the inf,

$$\mu^*(A) \leq \inf \left\{ \sum_{i=1}^{\infty} \rho(E_j) \right\} \subseteq \mu^*(B)$$

(Subadditivity) Take $\bigcup_{j=1}^{\infty} A_j$ for all $A_j \in P(X)$.

By definition of inf, $\forall \varepsilon > 0$ there exists $E_{jk} \subseteq \mathcal{E}$ such that

$$\sum_{i=1}^{\infty} \rho(E_{jk}) \leq \mu^*(A_j) + \frac{\varepsilon}{2^j}$$

so

$$\bigcup_{j=1}^{\infty} A_j \subseteq \bigcup_{jk}^{\infty} E_{jk}$$

for $E_{jk} \in \mathcal{E}$.

Then

$$\begin{aligned}
\mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) &\leq \sum_{j,k}^{\infty} \rho(E_{jk}) \\
&\leq \sum_{j=1}^{\infty} \mu^*(A_j) + \varepsilon
\end{aligned}$$

Then certainly,

$$\mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} \mu^*(A_j)$$

μ^*-measurable (Carathéodory Criterion): a collection $\mathcal{M}$ of subsets of X is μ^*-measurable if, given $A \in \mathcal{M}$, for all $E \subseteq P(X)$,

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

And in fact, it suffices to show

$$\mu^*(E) \geq \mu^*(A \cup E) + \mu^*(E \cap A^c)$$

2.4 Oct 01

Carathéodory Procedure: If μ^* is an outer measure and $\mathcal{M}$ are μ^*-measurable sets, then $\mathcal{M}$ is a σ-algebra and $\mu^*\big|_{\mathcal{M}}$ is a measure on $\mathcal{M}$

Proof:

STEP 1. $A \in \mathcal{M}$, $A^c \in \mathcal{M}$ by definition

---

STEP 2. Let $A, B \in \mathcal{M}$ and $A \cup B \in \mathcal{M}$. Then

$$\mu^*(A \cup B) = \mu^*(A) + \mu^*(B),\ A \cap B = \emptyset$$

Since $\mu^* < \infty$ by definition, it suffices to show

$$\begin{aligned}
\mu^*(E) &\geq \mu^*(E \cap A) + \mu^*(E \cap A^c) \\
&\geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)
\end{aligned}$$

Since $E \in \mathcal{M}$ satisfies the Carathéodory Criterion (by assumption),

$$\begin{aligned}
\mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) \\
&= \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c)
\end{aligned}$$

Now consider $A \cup B$. Using set algebra,

$$A \cup B = (A \cap B) \cup (A \cap B^c) \cup (B \cap A^c)$$

(and this matches the first three terms above very nicely)

Then

$$E \cap (A \cup B) = (E \cap A \cap B) \cup (E \cap A \cap B^c) \cup (E \cap B \cap A^c)$$

By subadditivity of the outer measure,

$$\mu^*(E \cap (A \cup B)) \leq \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap B \cap A^c)$$

Further,

$$\begin{aligned} \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c) \\ \geq mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap B \cap A^c) \end{aligned}$$

which is just

$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$$

Now take $\mu^*(A \cup B)$. Using the above,

$$\mu^*(A \cup B) = \mu^*(A \cup B \cap A) + \mu^*(A \cup B \cap A^c) = \mu^*(A) + \mu^*(B)$$

which is what we wanted to show.

Now we can inductively extend this pairwise additivity to a finite union.

Let $A_i \in \mathcal{M}$ and

$$\bigcup_{i=1}^{N} A_i \in \mathcal{M}$$

with $A_i \cap A_j = \emptyset$ for $i \neq j$

By induction,

$$\mu^* \left(\bigcup_{i=1}^{N} A_i \right) = \sum_{i=1}^{N} \mu^*(A_i)$$

STEP 3 (Countable Additivity): Let $A_i \in \mathcal{M}$ and $\bigcup_{i=1}^{\infty} A_i \in \mathcal{M}$ with $A_i \cap A_j = \emptyset$ for $i \neq j$.

Define

$$B_n = \bigcup_{j=1}^{n} A_j$$

Take a test set E with $\mu^*(E) < \infty$.

By induction on the Carathéodory Criterion,

$$\begin{aligned}
\mu^*(E) &= \mu^*(E \cap B_j) + \mu^*(E \cap B_j^c) \\
&= \mu^*\left(E \cap \bigcup_{j=1}^{n} A_j\right) + \mu^*\left(E \cap \bigcap_{j=1}^{n} A_j^c\right) \\
&= \sum_{j=1}^{n} \mu^*(E \cap A_j) + \mu^*(E \cap \bigcap_{j=1}^{n} A_j^c) && \text{Step 2} \\
&\geq \sum_{j=1}^{n} \mu^*(E \cap A_j) + \mu^*(E \cap \bigcap_{j=1}^{\infty} A_j^c) \\
&\geq \sum_{j=1}^{\infty} \mu^*(E \cap A_j) + \mu^*(E \cap \bigcap_{j=1}^{n} A_j^c) \\
&\geq \mu^*\left(E \cap \bigcup_{i=1}^{\infty} A_i\right) + \mu^*\left(E \cap \bigcap_{i=1}^{\infty} A_j^c\right) && \text{Subadditivity}
\end{aligned}$$

STEP 4 (Completeness) Let $\mu^*(A) = 0$, then $A \in \mathcal{M}$.

Take any $E \subseteq P(X)$. We want to show

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

But by monotonicity,

$$\mu^*(E \cap A) \leq \mu^*(A) = 0 \implies \mu^*(E \cap A) = 0$$

Using monotonicity again, we get the inequality. Hence, every set with outer-measure 0 is in $\mathcal{M}$.

Now take $A_1 \subseteq A$. By monotonicity, $\mu^*(A_1) = 0 \in \mathcal{M}$

Completeness: A measure space $(X, \mathcal{M}, \mu)$ is complete if $\forall A \in \mathcal{M}$ with $\mu(A) = 0$, then $B \in \mathcal{M}$ for all $B \subseteq A$.

Exercise: Use the Carathéodory Procedure to produce the Hausdorff Measure (HW 4)

Lebesgue Measure

On the real numbers, it would be very nice to have a measure μ such that $\mu((a, b)) = \rho(a, b) = b - a$.

Lemma: If $A \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i)$,

$$\mu^*(A) = \inf \sum_{n=1}^{\infty} \rho(a_i, b_i)$$

and $\mu^*((a, b)) = b - a$.

Then using the Carathéodory process, we get the Lebesgue Measure on $(\mathbb{R}, \mathcal{M}, \mu)$

Proposition (Faithfulness of the Lebesgue Measure): Let I be any interval (closed, open, half-open, etc.) on $\mathbb{R}$. Then $\mu(I) = \rho(I)$

Proof:

STEP 1. Suppose $I = [a, b]$ is closed and finite. $\forall \varepsilon > 0$, consider $(a - \varepsilon, b + \varepsilon) \supset [a, b]$.

By definition of inf,

$$\mu^*([a, b]) \leq \rho((a - \varepsilon, b + \varepsilon)) = b - a + 2\varepsilon$$

But by arbitrariness of ε, $\mu^*([a, b]) \leq b - a$

On the other hand, take $\bigcup_{i=1}^{\infty}(a_i, b_i) \supseteq [a, b]$. By Heine-Borel, there exists a finite cover for $[a, b]$ so we can take

$$[a, b] \subseteq \bigcup_{i=1}^{N}(a_i, b_i)$$

We want to show that

$$\sum_{i=1}^{N}(b_i - a_i) \geq b - a \implies \mu^*([a, b]) \geq b - a$$

a must be in some open interval, so call it $a \in (a_1, b_1)$. WLOG, suppose $b_1 \leq b$.

But then $b_1 \in (a_2, b_2)$. If $b_2 > b$, then $(a_1, b_1) \cup (a_2, b_2)$ would cover $[a, b]$ and

$$b_2 - a_2 + b_1 - a_1 \geq b - a$$

Therefore, assume $b_2 \leq b$. Inductively define $b_n \in (a_{n+1}, b_{n+1})$. But because we have a finite cover, this process is not infinite, i.e. for some N, $b_N > b$.

Then it suffices to show

$$b_N - a_N + b_{N-1} - a_{N-1} + \cdots + b_1 - a_1 \geq b - a$$

and by construction, each $-a_i + b_{i-1} > 0$

STEP 2. Now take any interval I. Consider

$$[a + \varepsilon, b - \varepsilon] \subseteq I \subseteq (a - \varepsilon, b + \varepsilon)$$

But this is an open cover so

$$\mu^*(I) \leq b - a + 2\varepsilon$$

But in part 1, we showed that

$$b - a - 2\varepsilon \leq \mu^*(I)$$

So by arbitrariness of ε, $\mu^*(I) = b - a$

2.5 Oct 03

Corollary: If A is a countable subset of $\mathbb{R}$, $\mu^*(A) = 0$. Furthermore, $[0,1]$ is not countable.

Proof: $\forall x \in \mathbb{R}$, $\{x\} = (x-\varepsilon, x+\varepsilon)$, so

$$\mu^*(\{x\}) \leq 2\varepsilon \implies \mu^*(\{x\}) = 0$$

Now suppose $A = \bigcup_{n=1}^\infty a_n$ with $a_n \in \mathbb{R}$.

By subadditivity,

$$\mu^*(A) \leq \sum_{n=1}^\infty \mu^*(\{a_n\}) = 0$$

Since $\mu^*([0,1]) = 1 \neq 0$, $[0,1]$ is not countable.

Proposition: $\mathcal{B}_\mathbb{R} \subseteq \mathcal{M}$

Proof: By the characterization of the Borel set on $\mathbb{R}$, it suffices to show that $(a,\infty) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

$\forall E \in \mathcal{M}$ with $\mu^*(E) < \infty$, we want to show

$$\begin{aligned}\mu(E) &\geq \mu^*(E \cap (a,\infty)) + \mu^*(E \cap (a,\infty)^c) \\ &= \mu^*(E \cap (a,\infty)) + \mu^*(E \cap (-\infty, a])\end{aligned}$$

For notational convenience, let

$$\begin{aligned}E_1 &= E \cap (a,\infty) \\ E_2 &= E \cap (-\infty, a]\end{aligned}$$

Let $\varepsilon > 0$. By the sharpness of the outer measure, $\exists \bigcup_{n=1}^\infty I_n$ such that

$$E \subseteq \bigcup_{n=1}^\infty I_n$$

and

$$\sum_{n=1}^\infty |I_n| < \mu^*(E) + \varepsilon$$

Then

$$\begin{aligned}E_1 &\subseteq \bigcup_{n=1}^\infty I_n \cap (a,\infty) \\ E_2 &\subseteq \bigcup_{n=1}^\infty I_n \cap (-\infty, a]\end{aligned}$$

so

$$\mu^*(E_1) \le \sum_{n=1}^{\infty} \mu^*(I_n \cap (a, \infty))$$
$$\mu^*(E_2) \le \sum_{n=1}^{\infty} \mu^*(I_n \cap (-\infty, a])$$

Now by the faithfulness of the Lebesgue Measure,

$$\mu(I_n) = \mu^*(I_n \cap (a, \infty)) + \mu^*(I_n \cap (-\infty, a])$$

so

$$\mu^*(E_1) + \mu^*(E_2) \le \sum_{n=1}^{\infty} \mu(I_n) \le \mu^*(E)$$

Transformations

Definitions: Given $E \subseteq \mathbb{R}$, we define

- (Translation) $E + a := \{x + a : x \in E\}$
- (Dilation) $rE := \{rx : x \in E\}$

Lemma: For the Lebesgue outer measure and $E \in \mathcal{M}$,

1. $\mu^*(E + a) = \mu^*(E)$
2. $\mu^*(rE) = |r| \, \mu^*(E)$

Proof Sketch: Let $E \subseteq \bigcup_{n=1}^{\infty} I_n$.

Certainly,

$$E + a \subseteq \bigcup_{n=1}^{\infty} \{I_n + a\}$$
$$rE \subseteq \bigcup_{n=1}^{\infty} \{|r| \, I_n\}$$

Then

$$\sum_{n=1}^{\infty} \rho(I_n) = \sum_{n=1}^{\infty} \rho(I_n + a) \ge \mu^*(E + a)$$

And taking the infimum,

$$\mu(E) \ge \mu^*(E + a)$$

For dilation, notice $|I_n| = \frac{1}{|r|} \, |rI_n|$. The result follows similarly.

The other direction is exactly the same.

Approximation of Measurable Sets

Lemma:

1. (Approximation from Above) $\forall E \subseteq P(X)$ and $\forall \varepsilon > 0$, then exists an open set O such that $E \subseteq O$ and
$$\mu(O) \geq \mu^*(E) \geq \mu(O) - \varepsilon$$
2. (Approximation from Below) $\forall E \subseteq \mathcal{M}$ and $\forall \varepsilon > 0$, $\exists K$ closed such that
$$\mu(K) \leq \mu(E) \leq \mu(K) + \varepsilon$$

Proof:

1. By definition of $\mu^*(E)$, $\exists O = \bigcup_{n=1}^{\infty} I_n \supset E$ such that
$$\mu^*(E) \geq \sum_{n=1}^{\infty} \rho(I_n) - \varepsilon$$
and by subadditivity, $\mu^*(E) \geq \mu^*(O) - \varepsilon$.

2. Assume $E \subseteq [a, b]$. Consider $E^c \cap [a, b]$. By part 1, $\exists O \supset E^c \cap [a, b]$ such that
$$\mu^*(E^c \cap [a, b]) \geq \mu^*(O) - \varepsilon$$
and with some algebra,
$$|b - a| - \mu^*(E^c) \leq |b - a| - \mu^*(O) + \varepsilon$$
By measurability,
$$\begin{aligned}\mu(E) &\leq |b - a| - \mu^*(O \cap [a, b]) - \mu^*(O \cap [a, b]^c) + \varepsilon \\ &\leq |b - a| - \mu^*(O \cap [a, b]) \\ &= \mu^*([a, b] \cap O^c) + \varepsilon\end{aligned}$$

Exercise: Complete the case $E \not\subseteq [a, b]$

Chapter 3

Measurable functions

3.1 Oct 08

Measurable function: Let $(X, \mathcal{M}, \mu)$ be a measure space. Let $f : X \to \mathbb{R}$. f is *measurable* iff $\forall \alpha \in \mathbb{R}$,

$$\{x \in X, f(x) > \alpha\} \in \mathcal{M}$$

Equivalently, f is measurable iff $\forall \alpha \in \mathbb{R}$, $\{f \geq \alpha\}$ is measurable.

Proposition: If f, g are measurable, so is

(a) $f + c$

(b) $f + g$

(c) fg

(d) cf

for $c \in \mathbb{R} \setminus \{0\}$

Proof: (a.) For all $\alpha \in \mathbb{R}$,

$$\{x : f(x) + c > \alpha\} = \{x : f(x) > \alpha - c\} \in \mathcal{M}$$

‘

(b.) Consider $\{x : f(x) + g(x) > \alpha\}$. We claim

$$\{x : f(x) + g(x) > \alpha\} = \bigcup_{q \in \mathbb{Q}} \{f(x) > \alpha - q\} \cap \{g(x) > q\}$$

Proof: Certainly,

$$\{x : f(x) + c > \alpha\} = \{x : g(x) > \alpha - f(x)\}$$

and for fixed x, we can invoke the density and countability of the rational numbers...

(c.) Consider $\{f^2 > \alpha\} = \{f > \sqrt{\alpha}\} \cup \{f < -\sqrt{\alpha}\}$. So f^2 is measurable. But then

$$fg = \frac{1}{2}\left[(f+g)^2 - (f-g)^2\right]$$

The rest follows from (b.)

(d.) Let $g(x) = c$ and the result follows from (c.)

Proposition: If $\{f_n\}$ are measurable, so are

(a) $\max_i f_i(x)$

(b) $\min_i f_i(x)$

(c) $\sup_n f_n$

(d) $\inf_n f_n$

(e) $\limsup_n f_n$

(f) $\liminf_n f_n$

Proof:

(a.) Suppose $n < \infty$. Then $\max_i f_i(x) > \alpha \iff \exists 1 \leq i_0 \leq n$ such that $f_{i_0}(x) > \alpha$.

Hence

$$\{x : \max_i f_i(x) > \alpha\} = \bigcup_{i=1}^{n} \{f_i(x) > \alpha\}$$

which are measurable by assumption.

(b.) Analogously,

$$\{x : \min_i f_i(x) > \alpha\} = \bigcap_{i=1}^{n} \{f_i(x) > \alpha\}$$

(c.) Now suppose we have countably many f_i. We claim

$$\{\sup_n f_n(x) > \alpha\} = \bigcup_{n=1}^{\infty} \{f_i(x) > \alpha\}$$

Proof: If $x \in \bigcup_{i=1}^{\infty} \{f_i(x) > \alpha\}$, then $\exists f_{i_0}(x) > \alpha$ so $\sup_n f_n(x) > \alpha$.

Now we want to show that if $\sup_n f_n(x) > \alpha$, $\exists i_0$ such that $f_{i_0}(x) > \alpha$. Suppose not.

Then $f_{i_0}(x) \leq \alpha \implies \sup_n f_n(x) \leq \alpha$, a contradiction.

(d.) Now we want to show $\inf_n f_n > \alpha$

Warning: when taking an infinite sequence, the strict inequality may not be preserved

Fact (Sup/Inf Parity):

$$\inf_n f_n = -\sup_n\{-f_n\}$$
$$\sup_n f_n = -\inf_n\{-f_n\}$$

and

$$\limsup_n f_n = -\liminf_n\{-f_n\}$$
$$\liminf_n f_n = -\limsup_n\{-f_n\}$$

Exercise: Prove the fact above

The proof of (d.) follows from (c.)

Preparations for Integration

Characteristic Function: Given $E \in \mathcal{M}$, we define

$$\chi_E = \mathbb{1}_E = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}$$

Homework: Show $\mathbb{1}_E$ is measurable

And now we can define an integral

$$\int_x \chi_E \, dx = \mu(E)$$

Simple functions: Let $E_1, \dots, E_n \in \mathcal{M}$. We define *simple functions*

$$\phi = a_1\chi_{E_1} + \cdots + a_n\chi_{E_n}$$

with *standard representation*

$$\phi = \sum_{i=1}^{n} a_i \chi_{E_i}$$

where $a_i < \infty$ and $a_i \neq a_j$ for $E_i \cap E_j = \emptyset$

Now we define a new integral

$$\int \phi := \sum_{i=1}^{n} a_i \mu(E_i)$$

Remark: The integral $\int_x \chi_E \, dx = \mu(E)$ corresponds to the Riemann integral and works by dividing the domain. $\int \phi$ is the *Lebesgue integration* and partitions the range of the function instead

Theorem: Let $(X, \mathcal{M})$ be a measurable space

(a) If $f : X \to [0, \infty]$ is measurable, then $\exists\{\phi_n\}$ of simple functions such that

$$0 \le \phi_1 \le \phi_2 \le \cdots \le f$$

Further, $\phi_n \to f$ pointwise on X and uniformly on any set which f is bounded.

(b) If $f : X \to \mathbb{R}$ is measurable, then $\exists\{\phi_n\}$ of simple functions such that

$$0 \le |\phi_1| \le |\phi_2| \le \cdots \le |f|$$

and $\phi_n \to f$ pointwise on X and uniformly on any set where f is bounded.

Proof:

(a) (Proof by construction) Fix n and choose $0 \le k \le 2^{2n} - 1$.

Define

$$\begin{aligned} E_n^k &= \{x : \frac{k}{2^n} < f(x) \le \frac{k+1}{2^n}, \quad 0 \le k \le 2^{2n} - 1\} \\ F_n &= \{x : f(x) \ge 2^n\} \end{aligned}$$

So we can choose

$$\phi_n(x) = \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} \chi_{E_n^k} + 2^n \chi_{F_n}$$

(Each n subdivides the range and then rounds down to each subdivision. As we iterate, each range is further divided in such a way that $\phi_{n+1} > \phi_n$ for all n because the value on E_{n+1} is rounded down to a higher value than E_n.)

So by construction,

$$0 \le \phi_n \le \phi_{n+1} \le f(x)$$

If we fix any $x < \infty$, for $f(x) < 2^{n_0}$ and $n \ge n_0$, we have

$$|\phi_n - f(x)| \le \left|(k+1)2^{-n} - k2^{-n}\right| = \frac{1}{2^n}$$

i.e., we have uniform convergence on any set where f is bounded.

(b) For a general function f, we can write $f = f^+ - f^-$ (so $|f| = f^+ + f^-$) where $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$. Since f^+ and f^- are nonnegative, we can apply part (a.) to each.

In particular,

$$0 \le |f_n| \le |f_{n+1}| \le |f(x)|$$

and the rest follows.

Conclusion: these simple functions can approximate *any* measurable function in this very strong sense.

3.2 Oct 10

Integration of Non-negative functions

In this section, assume $f \geq 0$.

For any $f : E_i \to [0, \infty]$, we consider the simple functions $\phi = \sum_{i=1}^{n} a_i - \chi_{E_i}$ in standard representation (a_i, a_j distinct when E_i, E_j disjoint)

Then, as above, the integral is defined as

$$\int \phi \, d\mu = \sum_{i=1}^{n} a_i \mu(E_i)$$

with respect to measure μ

Proposition: If ϕ and ψ are two simple functions,

1. If $c \geq 0$, $\int c\phi \, d\mu = c \int \phi \, d\mu$
2. $\int \phi + \psi \, d\mu = \int \phi \, d\mu + \int \psi \, d\mu$
3. If $\phi \leq \psi$, then $\int \phi \, d\mu \leq \int \psi \, d\mu$

Proof:

1) is trivial

2) Let $\phi = \sum_{i=1}^{n} a_i \chi_{E_i}$ and $\psi = \sum_{j=1}^{m} b_j \chi_{F_j}$. Then

$$\begin{aligned}
\int \phi \, d\mu + \int \psi \, d\mu &= \sum_{i=1}^{n} a_i \mu(E_i) + \sum_{j=1}^{m} b_j \mu(F_j) \\
&= \sum_{i=1}^{n} a_i \mu(E_i \cap (\bigcup F_j)) + \sum_{j=1}^{m} b_j \mu(F_j \cap E_i) \\
&= \sum_{i=1}^{n} a_i \mu(E_i \cap F_j) + \sum_{j=1}^{m} b_j \mu(E_i \cap F_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} (a_i + b_j) \mu(E_i \cap F_j)
\end{aligned}$$

Now we can cut $\phi = \sum_{i=1}^{n} a_i \chi_{E_i \cap \bigcup_{j=1}^{n} F_j}$ and ψ similarly.

Lemma: $\chi_E + \chi_F = \chi_{E \cup F}$

Proof: Exercise

So

$$\begin{aligned}\psi + \phi &= \sum_{i,j} a_i \chi_{E_i \cap F_j} + \sum_{i,j} b_j \chi_{E_i \cap F_j} \\ &= \sum_{i,j} (a_i + b_j) \chi_{E_i \cap F_j}\end{aligned}$$

However, integration is only defined in standard representation. But E_i, E_j are pairwise disjoint so $(a_i + b_i)$ distinct.

By definition,

$$\psi + \phi = \sum (a_i + b_j) \mu(E_i \cap F_j) = RHS$$

3) **Claim:** $\phi \leq \psi$ iff $a_i \leq b_i$ on $\chi_{E_i \cap E_j}$

Exercise

The result follows.

Integrals on subsets: Let $A \in \mathcal{M}$. Then

$$\int_A \phi \, d\mu = \int_X \chi_A \cdot \phi \, d\mu$$

with $\chi_{A_1} \cdot \chi_{A_2} = \chi_{A_1 \cap A_2}$

Proposition: For $A \in \mathcal{M}$, the mapping $A \to \int_A \phi \, d\mu$ is a measure on $\mathcal{M}$

Proof:

The hardest part is to show that $\int_A \phi \, d\mu$ is countably additive, that is

$$\int_{\bigcup A_i} \phi = \sum_{i=1}^{\infty} \int_{A_i} \phi$$

By definition,

$$\begin{aligned}\int_{\bigcup A_i} \phi &= \int \chi_{\bigcup A_i} \cdot \phi \\ &= \sum_{i=1}^{n} a_i \mu(E_i \cap \bigcup_{j=1}^{\infty} A_j)\end{aligned}$$

So it suffices to show that

$$\sum_{i=1}^{n} a_i \mu(E_i \cap \bigcup_{j=1}^{\infty} A_j) = \sum a_n \left[\sum_{i=1,k=1}^{\infty} \mu(E_i \cap A_j) \right]$$

and in fact all we need is

$$\mu(E_i \cap \bigcup_{j=1}^{\infty} A_j) = \sum_{ij} \mu(E_i \cap A_j)$$

which follows from countable additivity of the measure.

Integral of general functions: For a general function $f \geq 0$ on $(X, \mathcal{M}, \mu)$, we define

$$\int f \, d\mu = \sup_{\phi \leq f} \int \phi \, d\mu$$

for all simple functions ϕ

Properties: For $f \geq 0$, we have the following properties:

1. $c \int f \, d\mu = \int cf \, d\mu$
2. If $f \leq g$, then $\int f \, d\mu \leq \int g \, d\mu$
3. $\int f + g \, d\mu = \int f \, d\mu + \int g \, d\mu$

Proof:

1. We have a 1-1 correspondence between cf and f so the result follows from the same proof as above

2. Follows from sup and definition

3. The third one is much harder and we will need to come back to it.

Theorem (Monotone Convergence Theorem): Assume $0 \leq f_n \leq f_{n+1}$ for f measurable. Then

$$\int \lim_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu$$

Proof:

$\lim_{n\to\infty} f_n$ is increasing so we can say $\lim_{n\to\infty} f_n = f$ in the extended reals and the LHS is well defined.

From monotonicity,

$$\lim_{n\to\infty} \int f_n \leq \int f$$

The other direction is harder.

Take any simple function $\phi \leq f$ and pick $0 < \alpha < 1$. Consider

$$E_n = \{x : f_n(x) \geq \alpha\phi, \ \phi > 0\}$$

for fixed x.

Since f_n is non-decreasing, $E_n \subseteq E_{n+1}$. We claim

$$\bigcup_{n=1}^{\infty} E_n = \{f(x) > \alpha\phi(x), \ \phi > 0\} := E$$

$\subseteq$ is clear. For $\supseteq$, notice $\forall x, \ f_n \to f > \alpha\phi(x) \implies \exists n_0$ such that for $n \geq n_0, \ f_n(x) > \alpha\phi(x)$.

Now take

$$\int f_n \geq \int f_n \chi_{E_n} = \int_{E_n} f_n > \alpha \int_{E_n} \phi$$

By the proposition above, we can take $\nu(E_n) = \int_{E_n} \phi$ to be a measure so

$$\alpha \int_{E_n} \phi = \alpha\nu(E_n) \overset{n\to\infty}{\longrightarrow} \alpha\nu\left(\bigcup_{n=1}^{\infty} E_n\right)$$

Taking the limit, we cannot keep the strict inequality, but we do have

$$\int f_n \geq \int f_n \,\chi_{E_n} \geq \alpha\nu\left(\bigcup_{n=1}^{\infty} E_n\right) \alpha\nu(E)$$

and the integral of E is just

$$\alpha\nu(E) = \alpha \int \phi\chi_{\{f(x)>\alpha\phi(x),\, \phi>0\}}$$

Since $\phi \leq f$, for any x, $\phi(x) > 0$ so $\alpha\phi(x) < f(x)$. Note that for $\phi(x) = 0$, $\int_{\phi=0} \phi = 0$ so

$$\alpha \int \phi\chi_{\{f(x)>\alpha\phi(x),\, \phi>0\}} = \alpha \int \phi$$

Exercise: Check $\int \phi\chi_{\{f(x)>\alpha\phi(x),\, \phi>0\}} = \alpha \int \phi$ using the standard expression

But by definition,

$$\int f \geq \sup_{\phi\leq f} \int \phi = \int f$$

Fatou's Lemma: Let $0 \leq f_n$. Then

$$\int \liminf f_n \, d\mu \leq \liminf \int f_n \, d\mu$$

Proof: By definition,

$$\int \liminf f_n = \sup_n \inf_{k\geq n} f_k$$

Note $\inf_{k\geq n} f_k$ is increasing in n.

We claim

$$\sup_n \inf_{k\geq n} f_k = \lim_{n\to\infty} \inf_{k\geq n} f_k$$

Exercise: Check the above claim

And then by the Monotone Convergence Theorem,

$$\int \lim f_n = \int \lim_{n\to\infty} \inf_{k\geq n} f_k = \lim_{n\to\infty} \int \inf_{k\geq n} f_k$$

For any $m \geq n$,

$$\int \inf_{k\geq n} f_k \leq \int f_m$$

but this is true for all n so

$$\int \inf_{k \geq n} f_k \leq \inf_{k \geq n} \int f_k$$

Taking limits,

$$\begin{aligned} \lim_{n \to \infty} \int \inf_{k \geq n} f_k &\leq \lim_{n \to \infty} \inf_{k \geq n} \int f_k \\ &= \liminf \int f_k \end{aligned}$$

3.3 Oct 17

Corollary of MCT: If $f_n \geq 0$,

$$\sum_{n=1}^{\infty} \int f_n \, d\mu = \int \sum_{n=1}^{\infty} f_n \, d\mu$$

Proof: Define $F_N = \sum_{n=1}^{N} f_n$. Clearly F_N is increasing so by MCT,

$$\lim_{N \to \infty} \int F_N \, d\mu = \int \lim_{N \to \infty} F_N \, d\mu$$

and we are done.

Now at last we can verify the that the integral is additive for general functions by just taking the finite sum and applying this corollary.

Theorem (Lebesgue Dominated Convergence): Assume $f_n \geq 0$, $f_n \leq g$ for g fixed function with $\int g \, d\mu < \infty$. If $\lim_{n \to \infty} f_n = f$, then

$$\lim_{n \to \infty} \int f_n \, d\mu = \int f \, d\mu$$

Proof: By Fatou,

$$\int \liminf f_n \, d\mu \leq \lim \int f_n \, d\mu$$

but $\liminf f_n = f$ so

$$\int f \, d\mu \leq \liminf \int f_n \, d\mu$$

So it suffices to show

$$\limsup \int f_n \, d\mu = \liminf \int f_n = \int f$$

Further, $g - f_n \geq 0$, so again by Fatou

$$\int \liminf \, (g - f_n) \, d\mu \leq \liminf \int (g - f_n) \, d\mu$$

but $\lim f_n = f$ so

$$\int \liminf\, (g - f_n)\; d\mu = \int (g - f)\; d\mu$$

On the RHS,

$$\begin{aligned}\liminf \int (g - f_n)\; d\mu &= \liminf \left[\int g\; d\mu - \int f_n\; d\mu\right] \\ &\overset{*}{=} \int g\; d\mu + \liminf \left[-\int f_n\; d\mu\right]\end{aligned}$$

Exercise: Verify (*) above

Now using the parity of the $\liminf$,

$$RHS = \int g\; d\mu - \limsup \int f_n\; d\mu$$

Then,

$$\begin{gathered}\int g - \int f \leq \int g - \limsup \int f_n \\ \int f \geq \limsup \int f_n\end{gathered}$$

Warning: that subtraction only works because $\int g$ is finite by assumption

All together,

$$\limsup \int f_n\; d\mu \leq \int f\; d\mu \leq \limsup \int f_n\; d\mu$$

Integrable function: f is *integrable* if $\int_X f\; d\mu < \infty$

Zero Measure Sets

Almost Everywhere: We say a statement S is valid *almost everywhere* if $\exists E$, $\mu(E) = 0$ such that S is valid on $X \setminus E$

Lemma:

1. If $\mu(E) = 0$, then $\int_E f = 0$ for any $f \geq 0$
2. If $\int_X f\; d\mu = 0$ for $f \geq 0$, then $f = 0$ almost everywhere
3. If $\int_X f\; d\mu < \infty$, then $f < \infty$ almost everywhere

Proof:

1. Choose any simple function $\phi \leq f$. Notice

$$\begin{aligned}
\int_E \phi \, d\mu &= \int_X \mathbb{1}_E \cdot \phi \, d\mu \\
&= \int_X \mathbb{1}_E \sum_{i=1}^m a_i \mathbb{1}_{E_i} \, d\mu \\
&= \int_X \sum_{i=1}^m a_i \mathbb{1} E_i \cap E \, d\mu \\
&= \sum_{i=1}^m a_i \mu(E_i \cap E) \\
&\leq \mu(E) = 0
\end{aligned}$$

So

$$\int_E f = \sup_{\phi \leq f} \int_E \phi = 0$$

for ϕ simple function.

Remark: This depends on our convention that $0 \cdot \infty = 0$

2. Consider $E = \{x : f(x) > 0\}$. It suffices to show $\mu(E) = 0$ (because $E^c = \{x : f = 0\}$).

We can write

$$E = \bigcup_{n=1}^{\infty} \{x : f(x) \geq \frac{1}{n}\} = \bigcup_{n=1}^{\infty} E_n$$

so

$$\int_X f \, d\mu \geq \int_X \mathbb{1}_{E_n} \, d\mu = \int_{E_n} f \, d\mu$$

and by monotonicity,

$$\int_{E_n} f \, d\mu \geq \int_{E_n} \frac{1}{n} d\mu = \frac{1}{n} \int_{E_n} \, d\mu = \frac{1}{n} \mu(E_n)$$

But $\int_X f \, d\mu = 0$ so $\mu(E_n) = 0$ for all n and

$$\mu(E) = \lim_{n \to \infty} \mu(E_n) = 0$$

3. This statement is equivalent to the claim that $F = \{f(x) : f(x) = \infty\}$ has measure zero.

Let $F_n = \{x : f(x) \geq n\}$ so $F \subseteq F_n$ for all n, and in fact, $F = \bigcap n = 1^{\infty} F_n$

By exactly the same argument as before,

$$\begin{aligned}
\int_X f \, d\mu &\geq \int_X \mathbb{1}_{F_n} f \, d\mu \\
&= \int_{F_n} f \, d\mu \\
&\geq n \mu(F_n)
\end{aligned}$$

so

$$\mu(F_n) = \lim_{n \to \infty} \mu(F_n) = 0$$

Note that this gives a powerful fact: for all $\int f < \infty$,

$$\int f \, d\mu = \int_{F^c} f \, d\mu = \int_{\{x : f < \infty\}} f \, d\mu$$

Let $G = \{x : f < \infty\}$. we know

$$G = \bigcup_{n=1}^{\infty} \{x : f(x) \leq n\} = \bigcup_{n=1}^{\infty} G_n$$

and $G_n \uparrow G$ so $f \mathbb{1}_{G_n} \to f \mathbb{1}_G$ (pointwise) and $\mathbb{1}_{G_n} \leq \mathbb{1}_{G_{n+1}}$ so

$$\int_X f \, d\mu = \lim_{n \to \infty} \int f \mathbb{1}_{G_n} \, d\mu = \lim_{n \to \infty} \int_{\{x : f \leq n\}} f \, d\mu$$

Exercise: Prove that for MCT, Fatou, and LDC it suffices to take the condition to be true almost everywhere.

For example, the MCT holds if $f_n \leq f_{n+1}$ almost everywhere.

Integration for General Functions

For any function f, we can write $f = f^+ - f^-$ where $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$. This is powerful because both f^+ and f^- are nonnegative and measurable.

$\mathcal{L}^1$ **Space:** We say $f \in \mathcal{L}^1$ (i.e. f is integrable) if

$$\int |f| \, d\mu < \infty$$

or equivalently,

$$\int f^+ \, d\mu + \int f^- \, d\mu < \infty$$

In this case, we define

$$\int d \, d\mu = \int (f^+ f^-) \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu$$

Proposition: If $f, g \in \mathcal{L}^1$ and $a, b \in \mathbb{R}$, then

1. $af + bg \in \mathcal{L}^1$
2. $\int af + bg \, d\mu = a \int f \, d\mu + b \int g \, d\mu$

Proof: By the triangle inequality

$$|af + bg| \leq |a| \; |f| + |b| \; |g|$$

so

$$\int |af + bg| \; d\mu \leq |a| \int |f| \; d\mu + |b| \int |g| \; d\mu < \infty$$

First let us check the easier linearity condition $\int f + g \, d\mu = \int f \, d\mu + \int g \, d\mu$.

Define $h = f + g$ so $f^+ - f^- + g^+ - g^- = h^+ - h^-$. We would like to rearrange this so we have two nonnegative functions. But we need to be careful that each of these is finite. And in fact, this is true a.e. since f and g are integrable.

So

$$f^+ g^+ + h^- = h^+ + f^- + g^- \geq 0$$

which means we can integrate. Let $X \setminus E$ be the set where all six terms are finite. Then

$$\int_{X \setminus E} h \, d\mu = \int_{X \setminus E} f \, d\mu + \int_{X \setminus E} g \, d\mu$$

3.4 Oct 22

Complex-Valued Functions

$$\int f = \int \operatorname{Re}(F) + i \int \operatorname{Im}(F)$$

Sign Functions: Define

- $\operatorname{sign} x = x / |x|$ for $x \in \mathbb{R} \setminus \{0\}$
- $\operatorname{sign} z = z / |z|$ for $z \in \mathbb{C} \setminus \{0\}$

so we have the useful property $\overline{\operatorname{sign} z}\, z = |z|$

Lemma: If $f \in \mathcal{L}^1$, then $\left| \int f \, d\mu \right| \leq \int |f| \, d\mu$

Proof: Assume $\int f \, d\mu \neq 0$. For f real,

$$\left| \int f \, d\mu \right| = \left| \int f^+ \, d\mu - \int f^- \, d\mu \right|$$

Since these integrals are real numbers,

$$\left| \int f^+ \, d\mu - \int f^- \, d\mu \right| \leq \left| \int f^+ \, d\mu + \int f^- \, d\mu \right| = \left| \int (f^+ + f^-) \, d\mu \right| = \int |f| \, d\mu$$

If f is complex, we want to normalize it to the real case. By the property above, we can write

$\alpha = \overline{\text{sign} \int f \, d\mu}$ so

$$\begin{aligned}
\left| \int f \, d\mu \right| &= \alpha \cdot \int f \, d\mu \\
&= \int \alpha f \, d\mu \\
&= \text{Re} \left(\int \alpha f \, d\mu \right) \\
&= \int \text{Re} \left(\alpha f \right) \, d\mu \\
&\leq \int \left| \alpha f \right| \, d\mu \\
&= \int \left| \alpha \right| \, \left| f \right| \, d\mu
\end{aligned}$$

Now we can invoke the LDC for non-negative functions ($f_n = f_n^+ - f_n^-$ with $|f_n^+| \leq |g| \implies \int f_n \, d\mu \to \int f \, d\mu$) and we are done.

Lemma: Assume $\sum_{n=1}^{\infty} \int |f_n| \, d\mu < \infty$. Then $\sum_{n=1}^{\infty} f_n$ converges and

$$\sum_{n=1}^{\infty} \int f_n \, d\mu = \int \sum_{n=1}^{\infty} f_n \, d\mu$$

Proof: Reduces to the positive case as in the proof above.

Approximation by $\mathcal{L}^1$ functions

Theorem: If $f \in L^1(d\mu)$, $\forall \varepsilon > 0$, there exists a simple function $\phi = \sum a_j \chi_j$ such that

$$\int |f - \phi| \; d\mu < \varepsilon$$

Proof: Recall the dyadic (partitioning the range) approximation $\phi_n \to f$ pointwise for $|\phi_n| \leq |f|$.

By LDC, $\int \phi_n \to \int f$ and

$$|\phi_n - f| \leq |\phi_n| + |f| \leq \int |f| \in \mathcal{L}^1$$

so

$$\lim_{n\to\infty} \int |\phi_n - f| \; d\mu = 0$$

Now $\forall \varepsilon > 0$,

$$\phi_n = \sum_{i=1}^{i_n} a_i E_i, \qquad E_i \cap E_j = \emptyset, \; a_i \neq 0$$

And if $\mu(E_j) < \infty$,

$$|a_j| \; \mu(E_j) = \int_{E_j} |\phi_n| \leq \int |f| \; d\mu < \infty$$

Theorem (Reduction to smooth functions in $\mathbb{R}$): Let μ be the Lebesgue-Stieltjes measure on $\mathbb{R}$. Then E_J (in the previous approximation) ϕ_n can be taken as a finite union of open intervals. Moreover, there exists a C^∞ function ϕ that vanishes outside a bounded interval such that

$$\int |f - \phi| \; d\mu < \varepsilon$$

Proof: $\mu(E_j)$ can be approximated by an open set so

$$\int O_j \mathbb{1} < \infty = \mu(O_j) = \sum_{k=1}^{\infty} \mu(I_k)$$

with I_k disjoint.

But since ϕ_n is approximated by finitely many open intervals,

$$\sum_{k=1}^{N} \mu(I_k) \to \mu(O_j) \to \mu(E_j)$$

But this is equivalent to

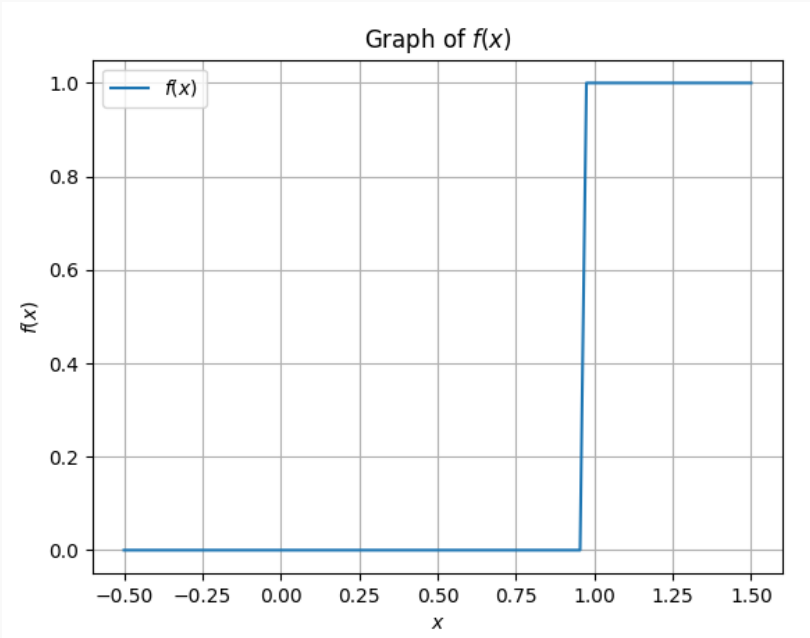
$$\sum_{k=1}^{N} \mu(I_k) = \sum_{I_k}^{N} \int_{I_k} \; d\mu \to \mu(E_j)$$

Therefore, it suffices to find a smooth function h such that

$$\int_{\mathbb{R}} \left| h - \mathbb{1}_{[a,b]} \right| < \varepsilon$$

Define the *One-sided Mollifier function*

$$f(x) = \begin{cases} \exp\left(-\frac{1}{e^{\frac{x}{1-x}}-1}\right) & 0 < x < 1 \\ 0 & x \leq 0 \\ 1 & x \geq 1 \end{cases}$$



And with a little analysis, we can show that this is smooth: $\partial^m f\Big|_{x=0,x=1} = 0$ for all $m \in \mathbb{N}$.

Define $h_\varepsilon(x) = f(\frac{x}{\varepsilon})$ so $0 \leq x/\varepsilon \leq 1$. Given $\mathbb{1}_{[a,b]}$,

$$g_\varepsilon(x) = \begin{cases} h(x-a-\varepsilon) & x \leq a+\varepsilon \\ 1 & a+\varepsilon \leq x \leq b \\ h_\varepsilon(b-x) & x \geq b \end{cases}$$

Clearly, $g_\varepsilon(x) \to \mathbb{1}_{[a,b]}$ as $\varepsilon \to 0$.

And further, $0 \leq g_\varepsilon(x) \leq \mathbb{1}_{[a-1,b+1]}$ for $\varepsilon << 1$ so by the LDC,

$$\int g_\varepsilon(x)\, d\mu \to \int \mathbb{1}_{[a,b]}\, d\mu$$

or

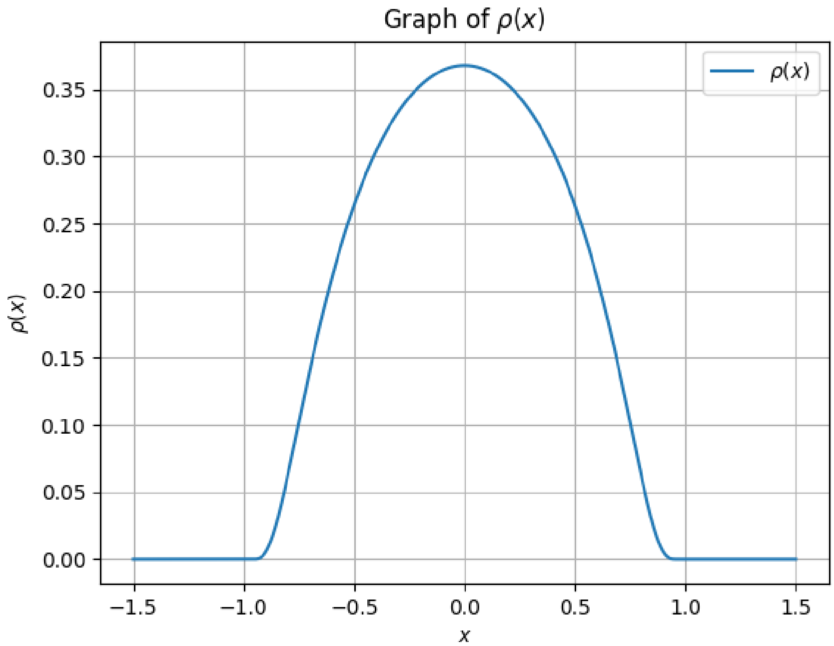
$$\int \left|g_\varepsilon - \mathbb{1}_{[a,b]}\right|\, d\mu \to \varepsilon$$

Remark: we could also define a (more standard) *Symmetric Mollifier function*

$$\rho(x) = \begin{cases} \exp(-\frac{1}{1-|x|^2}) & |x| < 1 \\ 0 & |x| > 1 \end{cases}$$

so

$$j = \frac{\rho}{\int \rho\, d\mu} \int j\, d\mu = 1$$



3.5 Oct 24

Theorem (Integrals with Parameter): Let $f(x,t) : X \times [a,b) \to \mathbb{C}$ where $f(;,t) : X \to \mathbb{C}$ is integrable for $t \in [a,b]$. Define

$$F(t) = \int f(x,t)\, d\mu$$

(a) Suppose $\exists g \in \mathcal{L}^1$ such that $|f(x,t)| \leq g(x)$ (independent of t) a.e. t. If $\lim_{t\to t_0} f(x,t) = f(x,t_0)$, for a.e. t, then

$$\lim_{t\to t_0} F(t) = F(t_0)$$

In particular, if $f(x,t)$ is continuous in t, then $F(t)$ is continuous.

(b) Assume $\frac{\partial f}{\partial t}$ exists and $\exists g \in L^1_\mu$ such that

$$\left|\frac{\partial f}{\partial t}(x,t)\right| \leq g(x)$$

then

$$F'(t) = \int \frac{\partial f}{\partial t}(x,t)\, d\mu$$

Proof:

(a) Choose any $t_n \to t_0$. Define $f_n(x) = f(x,t_n)$. By assumption $|f_n(x)| = |f(x,t_n)| \leq g \in \mathcal{L}^1$. By LDC,

$$\lim_{n\to\infty} \int f_n(x)\, d\mu = \int \lim_{n\to\infty} f_n(x)\, d\mu = \int f(x,t_0)\, d\mu$$

But t_n is arbitrary so

$$\lim_{t\to t_0} \int f(x,t)\, d\mu = \int f(x,t_0)\, d\mu$$

(b) Choose

$$h_n(x) = \frac{f(x,t_n) - f(x,t_0)}{t_n - t_0}$$

so for ξ between t_n and t_0,

$$|h_n(x)| = \left|\frac{\partial f}{\partial t}(x,\xi_n)\right| \le g \in L^1$$

Once again by LDC,

$$\lim_{n\to\infty} \int h_n(x)\, d\mu = \int \lim h_n(x)\, d\mu = \int \partial_t f(x,t_0)\, d\mu$$

Riemann Integrals

Recall: For $f : [a,b] \to \mathbb{R}$, we define the Riemann integral $(R)\int_a^b f\, dx$ by

$$\begin{aligned}
S_p(f) &= \sum_{j=1}^{n} M_j(t_j - t_{j-1}) \\
s_p(f) &= \sum_{j=1}^{n} m_j(t_j - t_{j-1}) \\
M_j &= \sup_{[t_{j-1},t_j]} f(x) \qquad\qquad m_j = \inf_{[t_{j-1},t_j]} f(x) \\
\overline{I}_a^b &= \inf_P S_p f, \quad \underline{I}_a^b = \sup_P s_p f
\end{aligned}$$

where

$$(R)\int_a^b f\, dx = \overline{I}_a^b = \underline{I}_a^b$$

Theorem: Let f be a bounded real function and μ the Lebesgue measure on $\mathbb{R}$.

(a) If f is Riemann integrable, then f is Lebesgue integrable and $(R)\int_a^b f(x)\, dx = \int_a^b f\, d\mu$

(b) f is Riemann integrable $\iff$ $\{x : f(x) \text{ not continuous on } [a,b]\}$ has Lebesgue measure zero

Proof:

(a) Notice that $\mathbb{1}_{[t_j,t_{j+1}]}$ is building block of Riemann integration and $\mathbb{1}_{E_j=\{j<f\le j+1\}}$ (in general *very* complicated) is the building block of Lebesgue integration.

Assume f is R-integrable. Then, $\exists P_k, P_m$ such that

$$\begin{aligned}
G_{P_k} &= \sum_{j=1}^{n} M_j \mathbb{1}_{(t_{j-1},t_j]} \implies \int G_{P_k} = \sum_{j=1}^{n} M_j(t_j - t_{j-1}) = S_{P_k}(f) \\
g_{P_m} &= \sum_{j=1}^{m} m_j \mathbb{1}_{(t_{j-1},t_j]} \implies \int g_{P,m} \sum_{j=1}^{m} m_j(t_j - t_{j-1}) = s_{P_m}(f)
\end{aligned}$$

so

$$\lim_{n\to\infty}\int G_{P_n} = (R)\int_a^b f$$
$$\lim_{m\to\infty} s_{P_m}(f) = 0$$

Now if we take a refinement $P_n \subseteq P_{n+1}$, $G_{P_k} \downarrow G$ and $G_{P_m} \uparrow g$ (by def as sup and inf) so

$$(R)\int G_{P_n}\,dx \geq \int f\,d\mu \geq (R)\int g_{P_n}\,dx$$

Taking limits,

$$\lim\int G_{P_n}\,dx = (R)\int_a^b f\,dx$$

but $\int G_{P_n}\,dx = \int G_{P_n}\,d\mu$ so using the measure theory view and the LDC,

$$\lim\int G_{P_n}\,d\mu = \int G\,d\mu$$

Therefore,

$$(R)\int_a^b f \geq \int_a^b G\,d\mu \geq \int f\,d\mu \geq \int g\,d\mu = (R)\int d\,dx$$

So at last, $\int G\,d\mu = \int g\,d\mu \implies \int(G-g)\,d\mu \implies G = g$ a.e.

Homework: Prove part (b)

Modes of Convergence

Egorov Theorem: Suppose $\mu(X) < \infty$ and $f_n \to f$ a.e. Then $\forall\varepsilon > 0$, $\exists E \subseteq X$ such that $\mu(E) < \varepsilon$ with $f_n - f \to 0$ uniformly on E^c

Proof:

Suppose $f_n \to f$ everywhere.

Recall the set of no convergence:

$$\phi = \{x : \exists\delta_x > 0 \text{ s.t. } \forall N, \exists n \geq N\ \ |f_n(x) - f(x)| \geq \delta_x\}$$

We want to construct a uniform set to approximate ϕ.

For $k \in \mathbb{N}$, define

$$E_n(k) = \bigcup_{m=n}^{\infty}\{x : |f_m(x) - f(x)| \geq \frac{1}{k}\}$$

so

$$E_n^c(k) = \bigcap_{m=n}^{\infty} \{x : |f_m(x) - f(x)| \leq \frac{1}{k}\}$$

Clearly, for fixed k, $E_n(k) \downarrow$ in n. We claim

$$\bigcap_{n=1}^{\infty} E_n(k) = \emptyset$$

(because $\forall x \in \bigcap_{n=1}^{\infty} E_n(k),\ \forall n, \exists m \geq n$ s.t. $|f_m(x) - f(x)| \geq \frac{1}{k}$)

Since $\mu(X) < \infty$,

$$\lim_{n\to\infty} \mu(E_n(k)) = \mu(\phi)$$

and by assumption, $\mu(E_n(k)) = 0$.

Now $\forall \varepsilon > 0$, $\exists n_k < n_{k+1}$ s.t. $\mu(E_{n_k}(k)) < \frac{\varepsilon}{2^k}$.

Therefore,

$$\mu\left(\bigcup_{k=1}^{\infty} E_{n_k}(k)\right) < \sum_{k=1}^{\infty} \mu(E_{n_k}) \leq \varepsilon$$

We let $E = \bigcup_{k=1}^{\infty} E_{n_k}(k)$ and claim $f_n \to f$ uniformly in E^c.

We know $x \in E^c \iff x \in \bigcap_{k=1}^{\infty} E_{n_k}^c$ so by definition, $\forall k, \exists n_k$ s.t. $\forall m \geq n_k,\ |f_m(x) - f(x)| \leq \frac{1}{k}$.

Exercise: Generalize this proof to the case $f_n \to f$ a.e.

3.6 Oct 29

Modes of Convergence (Contd.)

Convergence in Measure: We say $f_n \to f$ *in measure* ($f_n \xrightarrow{\mu} f$) if $\forall \varepsilon > 0$,

$$\lim_{n\to\infty} \mu\{x : |f_n(x) - f(x)| \geq \varepsilon\} = 0$$

Convergence in $\mathcal{L}^1$: We say $f_n \to f$ in $\mathcal{L}^1$ if $\int |f_n - f|\ d\mu \to 0$.

Lemma: Assume $f_n \to f$ in $\mathcal{L}^1$. Then $f_n \xrightarrow{\mu} f$

Proof:

$$\lim \int |f_n - f|\ d\mu = 0$$

and

$$\begin{aligned}
\lim \int |f_n - f| \ d\mu &\geq \int |f_n - f| \mathbb{1}_{\{x:|f_n-f|\geq\varepsilon\}} \ d\mu && \text{(monotonicity)} \\
&= \int_{\{x:|f_n-f|\geq\varepsilon\}} |f_n - f| \ d\mu \\
&\geq \int_{\{x:|f_n-f|\geq\varepsilon\}} \varepsilon \ d\mu && \text{(monotonicity)} \\
&= \varepsilon\mu\{x : |f_n - f| \geq \varepsilon\}
\end{aligned}$$

for all $\varepsilon > 0$.

Letting $n \to \infty$,

$$0 = \lim \frac{1}{\varepsilon} \int |f_n - f| \ d\mu \geq \lim \mu\{x : |f_n - f| \geq \varepsilon\}$$

Remark: in the proof, we have the inequality

$$\int |f_n - f| \ d\mu \geq \varepsilon\mu(\{x : |f_n(x) - f(x)| \geq \varepsilon\})$$

this is the **Chebyshev Inequality**.

Cauchy in measure: We say f_n is *Cauchy w.r.t measure* if $\forall\varepsilon, \delta > 0$, $\exists N$ s.t. $\forall n, m \geq N$,

$$\mu\{x : |f_n(x) - f_m(x)| \geq \varepsilon\} < \delta$$

Theorem: Assume f_n is Cauchy w.r.t convergence in measure. Then $\exists f$ measurable and f_{n_j} a subsequence such that $f_{n_j} \to f$ a.e. and $f_n \xrightarrow{\mu} f$

Proof: Since f_n is Cauchy, $\forall j$, we may choose f_{n_j} s.t.

$$E_j = \{x : \left|f_{n_j}(x) - f_{n_{j+1}}(x)\right| \geq 2^{-j}\} \implies \mu(E_j) < 2^{-j}$$

STEP 1. $\mu(\limsup E_j) = 0$.

Note $\sum_{j=1}^{\infty} \mu(E_j) \leq \sum_{j=1}^{\infty} 2^{-j} < \infty$. But by a HW, we know precisely that $\mu(\limsup E_j) = 0$.

STEP 2. If $\forall x \notin \limsup E_j$, $f_{n_j}(x)$ is Cauchy.

Notice that $\forall x \notin \limsup E_j = \bigcap_{j=1}^{\infty} \bigcup_{m=j}^{\infty} E_m$, $x \in \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} E_j^c$. But this implies $\exists N$ s.t. $\forall n \geq N$,

$$x \in \bigcap_{j=n}^{\infty} E_j^c = \bigcap_{j=n}^{\infty} \{x : \left|f_{n_j}(x) - f_{n_{j+1}}(x)\right| < 2^{-j}\}$$

WTS $\exists j_0$ such that $j, l \geq j_0$ such that $\left|f_{n_j}(x) - f_{n_l}(x)\right| < \varepsilon$.

WLOG, assume $n_l \geq n_j$. Now,

$$\begin{aligned}
\left|f_{n_j}(x) - f_{n_l}(x)\right| &\leq \sum_{j \geq l} \left|f_{n_j}(x) - f_{n_{j+1}}(x)\right| + \left|f_{n_{j+1}}(x) - f_{n_{j+2}}(x)\right| + \cdots + \left|f_{n_{l-1}}(x) - f_{n_l}(x)\right| \\
&\leq 2^{-j} + 2^{-j-1} + \cdots + 2^{-j-l} \\
&\leq 2^{-j+1}
\end{aligned}$$

and hence it is Cauchy.

STEP 3. $f_n \xrightarrow{\mu} f$.

Consider $\{x : |f_n(x) - f(x)| \geq \varepsilon\}$. We claim

$$\{x : |f_n(x) - f(x)| \geq \varepsilon\} \subseteq \{x : \left|f_n(x) - f_{n_j}(x)\right| \geq \frac{\varepsilon}{2}\} \bigcup \{x : \left|f_{n_j}(x) - f(x)\right| \geq \frac{\varepsilon}{2}\}$$

Notice

$$\varepsilon \leq |f_n - f| \leq \left|f_n - f_{n_j}\right| + \left|f_{n_j} - f\right| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$$

So

$$\mu\{x : |f_n(x) - f(x)| \geq \varepsilon\} \leq \mu\{x : \left|f_n(x) - f_{n_j}(x)\right| \geq \frac{\varepsilon}{2}\} + \{x : \left|f_{n_j}(x) - f(x)\right| \geq \frac{\varepsilon}{2}\}$$

But by assumption (Cauchy), $\mu\{x : |f_n(x) - f(x)| \geq \frac{\varepsilon}{2}\} \to 0$.

Further,

$$\{x : \left|f_{n_j}(x) - f(x)\right| \geq \frac{\varepsilon}{2}\} = (\limsup E_j) \cup (\limsup E_j)^c$$

and for j sufficiently large,

$$\{\left|f_{n_j} - f_{n_{j+1}}\right| \geq 2^{-j}\} \leq 2^{-j+1}$$

so we are done.

Corollary 1: If $f_n \xrightarrow{\mu} f$, then $\exists n_j$ such that $f_{n_j} \to f$ a.e.

Corollary 2: If $f_n \to f \in \mathcal{L}^1$, $\exists n_j$ such that $f_{n_j} \to f$ a.e.

Examples:

- For $f_n = \frac{1}{n}\mathbb{1}_{[0,n]}$, $\int f_n \, d\mu = 1$ and $f_n \to 0$ uniformly but $f_n \not\to 0 \in \mathcal{L}^1$
- For $f_n = \mathbb{1}_{[n,n+1]}$, $\int f_n \, d\mu = 1$ and $f_n \to 0$ a.e. but

  $$\int |f_n - f| \, d\mu = \int |f_n| \, d\mu = 1 \neq 0$$

  so $f_n \not\to 0$ in $\mathcal{L}^1$. Do we at least have convergence in measure? No:

  $$\mu\{x : |f_n(x) - 0| > \frac{1}{2}\} = \mu(n, n+1) = 1$$

- $f_n = n\mathbb{1}_{[0,\frac{1}{n}]}$.

  As before, $\int f_n \, d\mu = 1$ and $\forall x_0 \neq 0$, $\exists n_{x_0}$ s.t. $\forall n \geq n_{x_0}$, $f_n(x_0) = 0$ so $f_n \to 0$ a.e. By the same argument, we can say $f_n \not\to f$ in $\mathcal{L}^1$.

  However, $\mu\{x : |f_n(x) - 0| > \varepsilon\} = \mu[0, \frac{1}{n}] = \frac{1}{n} \to 0$ so $f_n \xrightarrow{\mu} 0$.

Oct 31

Recall: last time, we showed that if $f_n \xrightarrow{\mu} f$, then $\exists n_j$ such that $f_{n_j} \to f$ a.e.

Proof Sketch: Suffices to show

$$\lim_{j\to\infty} \mu\{|f_{n_j} - f| \geq \frac{\varepsilon}{2}\} = 0$$

By construction, $\exists j_0$ such that $\forall j \geq j_0$,

$$\bigcap_{j\geq j_0}^{\infty} E_{n_j}^c \iff |f_{n_j} - f| \leq 2^{-j+1}$$

(strict inequality is not preserved in the limits)

But also,

$$\bigcup_{j\geq j_0}^{\infty} E_{n_j}^c \iff |f_{n_j} - f| \geq 2^{-j+1}$$

so

$$\mu\{|f_{n_j} - f| \geq \frac{\varepsilon}{2}\} \leq \mu\{|f_{n_j} - f| \geq 2^{j+1}\} \leq \sum_{j}^{\infty} \mu(E_{n_j}) \leq 2^{-j+1}$$

Remark: we have actually shown a stronger result: $f_{n_j} - f \to 0$ as $\mu(E_{n_j}) \to 0$. Technically, it suffices to show that they both go to 0 – not that it happens at the same time.

Example (Moving intervals)

Let $f_1 = \mathbb{1}_{[0,1]}$, $f_2 = \mathbb{1}_{[0,1/2]}$, $f_3 = \mathbb{1}_{[1/2,1]}$, $f_4 = \mathbb{1}_{[0,1/4]}$, $f_5 = \mathbb{1}_{[1/4,1/2]}$, $f_6 = \mathbb{1}_{[1/2,3/4]}$, $f_7 = \mathbb{1}_{[3/4,1]}$ etc.

Assume $n = 2^k + j$ with $j < 2^k$, then

$$f_n = \mathbb{1}_{[\frac{j}{2^k}, \frac{j+1}{2^k}]}$$

Note that the support of the intervals is shrinking to 0!

We claim:

1. $f_n \xrightarrow{\mu} 0$

   Proof: $\mu\{|f_n| > \varepsilon\} \leq \mu\{x : f_n(x) \neq 0\} \to 0$

2. $f_n \not\to 0$ a.e.

   Proof: $\forall x_0$, $\exists k$ such that $x_0 \in [\frac{j}{2^k}, \frac{j+1}{2^k}]$ for some j (nested interval). But then $f_n(x_0) = 1$ on a set with measure $2^{-k} > 0$.

Product Measures

Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two measure spaces. We want to construct a canonical product measure $\mu \times \nu$.

Intuitively, it would be very nice if for $A \in \mathcal{M}$, $B \in \mathcal{N}$, $\mu(A \times B) = \mu(A)\nu(B)$

Recall: We have earlier defined $\mathcal{M} \otimes \mathcal{N}$ to be the σ-algebra generated by the rectangles $A \times B$ with $A \in \mathcal{M}$, $B \in \mathbb{N}$.

Facts:

1. $A \times B \cap (E \times F) = (A \cap E) \times (B \cap F)$
2. $(A \times B)^c = (X \times B^c) \cup A^c \times B = (A \times B^c) \cup (A^c \times B^c) \cup (A^c \times B)$

Let $\mathcal{A}$ be the finite union of rectangles $A \times B$. By the fact, $\mathcal{A}$ is closed under (finite) complements, hence is an algebra.

Lemma: Assume for $A \in \mathcal{M}$, $B \in \mathcal{N}$ with $\mu(A), \mu(B) < \infty$,

$$A \times B = \bigcup_{i=1}^{\infty} A_i \times B_i$$

with $(A_i \times B_i) \cap (A_j \times B_j) = \emptyset$ for $i \neq j$. Then

$$\mu(A)\nu(B) = \sum_{i=1}^{\infty} \mu(A_i)\nu(B_i)$$

Proof:

By a property of the characteristic function,

$$\mathbb{1}_{A \times B}(x, y) = \mathbb{1}_A(x) \cdot \mathbb{1}_B(y)$$

We claim

$$\mathbb{1}_A(x) \cdot \mathbb{1}_B(y) = \sum_{i=1}^{\infty} \mathbb{1}_{A_i \times B_i}(x, y)$$

(*Proof:* Self evident by pairwise disjoint)

But then

$$\sum_{i=1}^{\infty} \mathbb{1}_{A_i \times B_i}(x, y) = \sum_{i=1}^{\infty} \mathbb{1}_{A_i}(x) \cdot \mathbb{1}_{B_i}(y)$$

Fix x and consider: as functions of y, the characteristic function is measurable. Hence, by the series version of MCT,

$$\mathbb{1}_A(x)\nu(B) = \sum_{i=1}^{\infty} \mathbb{1}_{A_i}(x) \cdot \nu(B_i)$$

But now this is just a function of x, so by the same argument,

$$\mu(A)\nu(B) = \sum_{i} \mu(A_i)\nu(B_i)$$

But this order was arbitrary! Instead fix y. By MCT with respect to x,

$$\mu(A)\mathbb{1}_B(y) = \sum_{i} \mu(A_i) \cdot \mathbb{1}_{B_i}(y)$$

and then integrating over y,

$$\mu(A)\nu(y) = \sum_{i} \mu(A_i)\nu(B_i)$$

which is exactly the same!

This is the essence of Fubini's Theorem.

Remark: the condition for this lemma is highly nontrivial. In general, it is extremely difficult to cover an arbitrary space in disjoint rectangles (see an earlier HW)

Now it remains to actually construct the product measure such that

$$\mu \times \nu(A \times B) = \mu(A)\nu(B)$$

Premeasure: we call μ_0 a *Premeasure* if for $A = \bigcup_{i=1}^{\infty} A_i$ with $\mu_0(A_i)$ well defined,

$$\mu_0(A) = \sum_{i=1}^{\infty} \mu_0(A_i)$$

Recall the outer measure

$$\mu^*(E) = \inf \sum_i \rho(A_i)$$

for $E \subseteq \bigcup_i A_i$ which immediately gives us a measure by Carathéodory.

Let's try to replicate this. Let $\mu(A \times B) = \mu(A)\nu(B)$ be set functions on $\mathcal{A}$, the set of finite disjoint unions of rectangles $A \times B$.

Define

$$\mu^*(E) = \inf \sum_i \mu(A_i)\nu(B_i)$$

for $E \subseteq \bigcup_i A_i \times B_i$. By results earlier in the class, this is an outer measure which gives us a measure $\mu \times \nu$.

Lemma:

1. $\mu \times \nu(A \times B) = \mu(A)\nu(B)$
2. $\mathcal{A}$ is $\mu \times \nu$-measurable

Proof:

1. Let $A \times B \subseteq \mathcal{A}$ and suppose $A \times B \subseteq \bigcup_{i=1}^{\infty} A_i \times B_i$.

Clearly,

$$A \times B = (A \times B) \cap \left(\bigcup_{i=1}^{\infty} A_i \times B_i \right) = \bigcup_{j=1}^{\infty} (A \times B \cap A_i \times B_i)$$

Let $B_n = E \cap \left(\bigcup_{i \in I} A_i \right)$

By the earlier lemma,

$$\sum_{i=1}^{\infty} \mu(A \cap A_i)\nu(B \cap B_i) = \mu(A)\nu(B)$$

Therefore,

$$(\mu \times \nu)(A \times B) \geq \mu(A)\nu(B)$$

But also, for any $A \times B$, take its trivial covering so

$$(\mu \times \nu)(A \times B) \leq \mu(A)\nu(B)$$

3.7 Nov 7

Recall: For measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$, we would like to construct a product measure $\mu \times \nu$ on $X \times Y$ such that

$$\mu \times \nu(A \times B) = \mu(A)\mu(B)$$

Last time, we showed that this is true on $\mathcal{A}$, a finite union of disjoint rectangles $A_i \times B_i$:

$$A \times B = \bigcup_{i=1}^{\infty} A_i \times B_i \implies \mu \times \nu(A \times B) = \sum_{i=1}^{\infty} \mu(A \cap A_i)\nu(B \cap B_i)$$

By our construction, the μ_0 that satisfies this property is a *premeasure* (i.e. it is countably additive but not necessarily defined on a σ-algebra). It would be very nice if we could extend this to all of $\mathcal{M} \otimes \mathcal{N}$ (just as we did for Outer Measures with Carathéodory).

And in fact, it is not too difficult. First recall that, in general,

$$\mu^*(E) = \inf_{\bigcup_i A \supseteq E} \sum_{j=1}^{\infty} \mu_0(A_j)$$

is an outer measure. We then invoke Carathéodory's Theorem to get a measure μ on $\mathcal{M} \otimes \mathcal{N}$ and we are done.

All that remains are a few specifics:

Proposition: If μ_0 is a premeasure on $\mathcal{A}$, then for μ^* defined by

$$\mu^*(E) = \inf_{\bigcup_i A \supseteq E} \sum_{j=1}^{\infty} \mu_0(A_j)$$

1. $\mu^*\big|_{\mathcal{A}} = \mu_0$
2. $\mathcal{A}$ is μ^*-measurable

Proof:

1. Take a set $E \subseteq A$. By choosing E as its own cover, $\mu^*(E) \leq \mu_0(E)$ by the inf.

Conversely, take any cover $\bigcup_{j=1}^{\infty} A_j \supseteq E$. WLOG, A_j are disjoint (or else we can repeat the argument with $B_j = \bigcup_{j=1}^{\infty} A_j \setminus A$).

Exercise: Verify the case where A_j are not disjoint.

Then,

$$\bigcup_{j=1}^{\infty} A_j \cap E = E \subseteq \mathcal{A} \implies \sum_{j=1}^{\infty} \mu(A_j \cap E) = \mu_0(E)$$

since μ_0 is a premeasure.

But we know also

$$\sum_{j=1}^{\infty} \mu_0(A_j) \geq \sum_{j=1}^{\infty} \mu_0(A_j \cap E)$$

so by def of inf,

$$\mu_0(E) \geq \inf \sum \mu_0(A_j) = \mu^*(E)$$

---

2. Choose a test set $E \subseteq P(X \times Y)$. Then, by definition, $\exists B_j^\infty \in \mathcal{A}$ such that for $E \subseteq \bigcup_j B_j$,

$$\mu^*(E) + \varepsilon > \sum_j \mu_0(B_j) = \sum_j \mu_0(B_j \cap A) + \mu_0(B_j \cap A^c)$$

(disjoint since $B_j \in \mathcal{A}$)

But now the first term gives us a cover for $E \cap A$ and the second term gives us a cover for $E \cap A^c$.

Certainly,

$$\sum_j \mu_0(B_j \cap A) + \mu_0(B_j \cap A^c) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

We now want two more powerful properties: uniqueness and faithfulness.

Theorem (Uniqueness of Extension): Let $\mathcal{A} \subseteq P(X)$ be an algebra, μ_0 a premeasure. Let $\mathcal{M}$ be the σ-algebra generated by $\mathcal{A}$ and μ the extension of μ_0 to $\mathcal{M}$. If ν is another measure on $\mathcal{M}$ that agrees with μ_0 on $\mathcal{A}$, then $\nu(E) \leq \mu(E)$ for all $E \in \mathcal{M}$ with $\mu(E) < \infty$

Proof:

STEP 1. Let $E \subseteq \mathcal{M}$ with $E \subseteq \bigcup_{j=1}^\infty A_j$ for $A_j \in \mathcal{A}$. Then,

$$\nu(E) \leq \sum_{j=1}^\infty \nu(A_j) = \sum_{j=1}^\infty \mu_0(A_j)$$

which implies $\nu(E)$ is a lower bound (by μ^* definition).

Hence, $\nu(E) \leq \mu(E)$.

STEP 2. $\mu(A) = \nu(A)$ for $A = \bigcup_{j=1}^\infty A_j$.

We note

$$\bigcup_{j=1}^\infty A_j \nearrow A$$

By taking the limit

$$\nu(A) = \lim_{n\to\infty} \nu(\bigcup_{j=1}^n A_j) = \lim_{n\to\infty} \mu(\bigcup_{j=1}^n A_j)$$

STEP 3. Take $A = \bigcup_{i=1}^\infty A_j$ such that $E \subseteq \bigcup_j A_j$ so

$$\mu(A) \leq \mu(E) + \varepsilon \implies \mu(A \setminus E) < \varepsilon$$

We claim $\mu(E) \leq \nu(E)$:

$$\begin{aligned}\mu(E) &\leq \mu(A) = \nu(A) \\ &= \nu(E) + \nu(A \setminus E) \\ &\leq \nu(E) + \mu(A \setminus E) \\ &< \nu(E) + \varepsilon \\ &\xrightarrow{\varepsilon \to 0} \nu(E)\end{aligned}$$

STEP 4.

Exercise: Prove that μ is σ-finite.

Corollary: Lebesgue Product Measure

Proof: HW (Construct a product measure via cubes).

By uniqueness, the measure from HW and the measure here are the same.

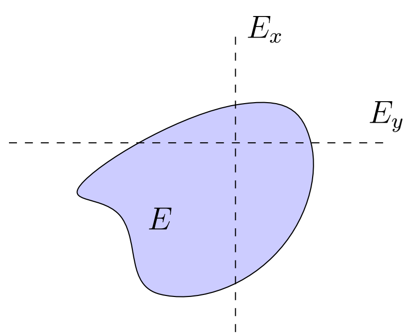
3.7.1 Fubini's Theorem

Consider $E \subseteq \mathcal{M} \otimes \mathcal{N}$. For all x, take sections of E:

$$E_x : \{y \in Y : (x, y) \in E\}$$

and similarly for y:

$$E_y : \{x \in X : (x, y) \in \mathbb{E}\}$$



Lemma:

1. $E \in \mathcal{M} \otimes \mathcal{N} \implies E_x \in \mathcal{N}$ and $E_y \in \mathcal{M}$
2. If $f(x, y)$ is a measurable function on $\mathcal{M} \otimes \mathcal{N}$, then $f_x(x, y)$ is measurable on y and $f^y(x, y)$ is measurable on x.

Proof:

1. Define $R = \{\forall x,\ E_x \in \mathcal{N}, \wedge \forall y,\ E_y \in \mathcal{M}\}$. It suffices to show that $R \supseteq \mathcal{M} \otimes \mathcal{N}$.

By the previous lemma, $\mathcal{A} \subseteq R$.

Exercise: Check that R is a σ-algebra. (Hint: $\mathcal{M}$ and $\mathcal{N}$ are σ-algebras)

But by definition, $\mathcal{M} \otimes \mathcal{N}$ is generated by $\mathcal{A}$ (hence the smallest σ-algebra containing $\mathcal{A}$), so $R \supseteq \mathcal{M} \otimes \mathcal{N}$.

2. $(f_x)^{-1}(B) = (f^{-1}(B))_x$ and $(f^y)^{-1}(A) = (f^{-1}(A))^y$ by properties of the preimage.

Exercise: Check that the preimage properties hold

By part 1, f_x, f_y are measurable.

Fubini-Tonelli Theorem: Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ-finite measure spaces. If $f \in \mathcal{L}^+(X \times Y)$, then $g(x) = \int f_x \, d\nu$ and $h(y) = \int f^y \, d\mu$ satisfies

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

Proof of Weak Form: By dyadic approximation, it suffices to consider $\mathbb{1}_E \approx \mathbb{1}_A$

STEP 1. If $\mu \times \nu(E) = 0$, then $\nu(E^x) = 0$ for almost every x WRT μ. Similarly, $\mu(E^y) = 0$ for almost every y WRT ν.

Since $\mu \times \nu(E) = 0$, choose $A_n \in \mathcal{A}$, $\mu \times \nu(A_n) \to 0$ so for $E \subseteq A_n$,

$$\mathbb{1}_E \leq \mathbb{1}_{A_n} \implies \mathbb{1}_{E_x} \leq \mathbb{1}_{A_{ny}} \implies \mathbb{1}_E(y) \leq \mathbb{1}_{A_n}(y)$$

Therefore,

$$\int \nu(E_x) = \int \int \mathbb{1}_{E_x} \, d\nu \, d\mu \leq \int \int \mathbb{1}_{A_{nx}} \, d\nu \, d\mu = \mu \times \nu(A_n) \overset{n \to \infty}{\longrightarrow} 0$$

and $\int \mu(E^y)$ follows similarly.

STEP 2. For $\mathbb{1}_E$, we have $\mathbb{1}_{A_n}$ such that $E \subseteq A_n$ and $\mu(E \setminus A_n) \to 0$. Hence, $\mathbb{1}_{A_n} \to \mathbb{1}_E$ in $\mathcal{L}^1(\mu \times \nu)$. By Chebyshev, it follows that $\mathbb{1}_{A_n} \to \mathbb{1}_E$ in measure. Hence, $\exists A_{n_k}$ such that $\mathbb{1}_{A_{n_k}} \to \mathbb{1}_E$ a.e. (WRT $\mu \times \nu$). So $\exists E^0$ with $\mu(E^0) = 0$ such that $\mathbb{1}_{A_{n_k}} \to \mathbb{1}_E$ on $E \setminus E^0$.

But then,

$$\int_{X \times Y \setminus E^0} \mathbb{1}_{A_n} \, d\mu \times \nu = \int_{X \times Y \setminus E^0} (\mathbb{1}_{A_n} \, d\mu) \, d\nu = \int_{X \times Y \setminus E^0} (\mathbb{1}_{A_n} \, d\nu) \, d\mu$$

(since E^0 has measure 0)

Taking $n \to \infty$, everything is bounded and in $\mathcal{L}^1$ so by LDC,

$$\int_{X \times Y \setminus E^0} \mathbb{1}_B ; d\mu \times \nu = \int_{X \times Y \setminus E^0} (\mathbb{1}_B \, d\mu) \, d\nu = \int_{X \times Y \setminus E^0} (\mathbb{1}_B \, d\nu) \, d\mu$$

and hence

$$\int_{X \times Y} \mathbb{1}_B ; d\mu \times \nu = \int_{X \times Y} (\mathbb{1}_B \, d\mu) \, d\nu = \int_{X \times Y} (\mathbb{1}_B \, d\nu) \, d\mu$$

STEP 3. Recall the dyadic approximation $0 \leq f \in \mathcal{L}^1(\mu \times \nu)$ and $f_n \nearrow f$ pointwise where f_n are simple functions.

Clearly, step 2 should hold for f_n. Hence,

$$\int \int f_n \, d\mu \, d\nu = \int \int f_n \, d\nu \, d\mu$$

We can also take sections $f_{n_x} \nearrow f_x$ and $f_n^y \nearrow f^y$ so taking the limit,

$$\int \int f \, d\mu \, d\nu = \int \int f \, d\nu \, d\mu$$

Remark: This works just fine for integration in general but we would still like a more general result that does not depend on losing a measure 0 set.

3.8 Nov 12

Last time, we gave a weak form of the proof of the Tonelli Theorem:

1. $(\mu \times \nu)E = 0 \implies \mu(E^y) = 0, \ \nu(E^x) = 0$ and measurable
2. By approximation of inf and A_n disjoint union,
$$||\mathbb{1}_{A_n} - \mathbb{1}_E||_{\mathcal{L}^1} \to 0 \implies \mathbb{1}_{A_{n_k}} \to \mathbb{1}_E \text{ a.e.}$$
3. Choose a zero measure set E_0 so $(\mu \times \nu)E_0 = 0$ and $\mathbb{1}_{A_{n_k}} \to \mathbb{1}_E$ on $E \setminus E_0$. Hence,
$$\mathbb{1}_{A_{n_k} \cap E_0^c} \to \mathbb{1}_{E \cap E_0^c}$$
pointwise.
4. Now,
$$\int \mathbb{1}_{A_{n_k}} \, d\mu \, d\nu = \int (\int \mathbb{1}_{A_{n_k}} \, d\mu) \, d\nu = \int (\int \mathbb{1}_{A_{n_k}} \, d\nu) \, d\mu$$
as $k \to \infty$,
$$\int \mathbb{1}_{E \setminus E_0} \, d\mu \, d\nu = \int (\int \mathbb{1}_{E \setminus E_0} \, d\mu) \, d\nu = \int (\int \mathbb{1}_{E \setminus E_0} \, d\nu) \, d\mu$$
Using LDC,
$$\int \mathbb{1}_E \, d\mu \, d\nu = \int \int \mathbb{1}_E \, d\mu \, d\nu = \int \int \mathbb{1}_E \, d\nu \, d\mu$$

But this still sacrifices a set of measure 0. We will offer a new proof that uses monotone convergence instead of convergence in norm.

But first, a lemma.

Monotone Class over X: $\mathcal{C} \subseteq P(x)$ is a *monotone class* if

- for $E_n \in \mathcal{C}$ and $E_n \nearrow$, then $\bigcup_{n=1}^{\infty} E_n \in \mathcal{C}$
- OR for $E_n \in \mathcal{C}$ and $E_n \searrow$, then $\bigcap_{n=1}^{\infty} E_n \in \mathcal{C}$

Lemma: Consider $A \subseteq P(X)$, an algebra. We consider the sigma algebra $\mathcal{M} = \mathcal{M}(A)$ and the monotone class $\mathcal{C} = \mathcal{C}(A)$ generated by A. Then $\mathcal{M} = \mathcal{C}$.

Proof: By definition, every σ-algebra is a monotone class so the smallest monotone class cannot be larger than the smallest σ-algebra: $\mathcal{C} \subseteq \mathcal{M}$.

Hence, it suffices to show that $\mathcal{C}$ is a σ-algebra.

STEP 1. Consider $\mathcal{C}(E) = \{F \subseteq \mathcal{C} : E \setminus F,\ F \setminus E,\ E \cap F \in \mathcal{C}\}$.

Clearly, $\emptyset \in \mathcal{C}(E)$ and $E \in \mathcal{C}(E)$.

Further, since we have the symmetric difference, $F \in \mathcal{C}(E) \iff E \in \mathcal{C}(F)$.

We also know that $\mathcal{C}(E)$ is a monotone class since if $F_n \nearrow F_\infty$, $E \setminus F \searrow$. Similarly, $F \setminus E \nearrow$ for $F_n \searrow F_\infty$.

STEP 2. For $E \subseteq A$, then $F \in \mathcal{C}(E)$ for all $F \in A$ (since A is an algebra). Therefore, $A \subseteq \mathcal{C}(E) \implies \mathcal{C}(A) \subseteq \mathcal{C}(E)$

STEP 3. $\mathcal{C}(A) \subseteq \mathcal{C}(F)$ for any $F \in \mathcal{C}$. Then $\forall F \in \mathcal{C}$, $F \in \mathcal{C}(E)$ for any $E \subseteq A$.

By symmetry, $E \subseteq \mathcal{C}(F)$. Hence, $\mathcal{C}(A) \subseteq \mathcal{C}(F)$ which implies $\mathcal{C}(A)$ is an algebra.

STEP 4.

Exercise: Show $\mathcal{C}(A)$ is a σ-algebra

Remark: This is incredibly powerful because it means we can always approximate a σ-algebra by monotone sets.

Tonelli Theorem:

$$\int \mathbb{1}_E \, d\mu \, d\nu = \int \left(\int \mathbb{1}_E \, d\mu \right) \, d\nu = \int \left(\int \mathbb{1}_E \, d\mu \right) \, d\nu$$

Strong form of the Proof: WLOG, assume μ, μ are finite (or else just approximate).

Denote $\mathcal{C} = \{\text{all sets } E \text{ s.t. the theorem holds}\}$.

Certainly, $A \subseteq \mathcal{C}$ (basic lemma).

We claim $\mathcal{C}$ is a monotone class, i.e. $E_n \nearrow E_\infty$ or $E_n \searrow E_\infty$ implies $E_\infty \in \mathcal{C}$. In fact, by construction, $\mathbb{1}_{E_n} \nearrow \mathbb{1}_{E_\infty}$ so taking sections,

$$\begin{cases} \mathbb{1}_{E_{n_x}} \nearrow \mathbb{1}_{E_{\infty x}} \quad \text{OR} \quad \mathbb{1}_{E_{n_x}} \searrow \mathbb{1}_{E_{\infty x}} \\ \mathbb{1}_{E_n^y} \nearrow \mathbb{1}_{E_\infty^y} \quad \text{OR} \quad \mathbb{1}_{E_n^y} \searrow \mathbb{1}_{E_\infty^y} \end{cases}$$

Now taking the limits (and using finite measure in the $\searrow$ case),

$$\int \mathbb{1}_{E_\infty} \, d\mu \, d\nu = \int \left(\int \mathbb{1}_{E_\infty} \, d\mu \right) \, d\nu = \int \left(\int \mathbb{1}_{E_\infty} \, d\mu \right) \, d\nu$$

so $E_\infty \subseteq \mathcal{C}$.

By the monotone class lemma, $\mathcal{M} = \mathcal{C}$.

3.9 Nov 14

3.9.1 The Lebesgue Measure on $\mathbb{R}^n$

Denote m the Lebesgue measure and $\mathcal{L}^n$ the class of Lebesgue measurable sets.

Theorem (Approximation): If $E \in \mathcal{L}^n$,

1. $m(E) = \inf\{m(U) : E \subseteq U,\ U \text{ open}\} = \sup\{m(K) : K \subseteq E,\ K \text{ compact}\}$
2. $E = A_1 \cup N_1 = A_2 \setminus N_2$, where A_1 is a F_σ set (i.e. union of open sets), A_2 is a G_δ set (i.e. intersection of closed sets), and $m(N_1) = m(N_2) = 0$. (In particular, A_1 and A_2 are Borel sets).
3. If $m(E) < \infty$, then $\forall \varepsilon > 0$, there exists a finite collection of disjoint rectangles such that

$$m\left(E \triangle \bigcup_{j=1}^{N} R_j\right) < \varepsilon$$

Proof Sketch:

1) $\forall \varepsilon > 0$, take a covering by disjoint rectangles $A_j \in \mathcal{A}$ such that $m(E) \leq \sum m(A_j) < m(E) + \varepsilon$. For each j, choose an open set U_j s.t. $m(U_j) \leq m(A_j) + \varepsilon 2^{-j}$ and $A_j \subseteq U_j$. Now

$$E \subseteq U = \bigcup_{j=1}^{\infty} U_j \implies \sum_{j=1}^{\infty} m(U_j) < m(E) + 2\varepsilon \implies m(E) = \inf\{m(U) : E \subseteq U\}$$

The proof for sup follows immediately from the 1-D proof earlier in the course.

2) Follows from Part 1 and the midterm

3) We may split A_j into very small sets δ so

$$m(E) \leq \sum_{1}^{N} m(A_j) + \sum_{N}^{\infty} m(\delta) < m(E) + \varepsilon$$

Choosing N large and using the finitude of $m(E)$,

$$m(E) - \varepsilon < \sum_{1}^{N} m(A_j) < m(E) + \varepsilon$$

Corollary of 3 (Lusin's Theorem): (globally defined) smooth functions approximate each $\mathbb{1}_{V_j}$:

$$\left\| \sum_{i=1}^{N} \mathbb{1}_{V_j} - \mathbb{1}_E \right\|_{L^1} < \varepsilon$$

Approximation for $\mathcal{L}^1$ functions (1D general): If $f \in \mathcal{L}^1(m)$, let $\varepsilon > 0$, then there exists a simple function $\phi = \sum_1^N a_i \mathbb{1}_{R_j}$ for R_j disjoint rectangles such that

$$\int |f - \phi| < \varepsilon$$

In particular, we can choose a continuous function f_c with compact support such that

$$\int |f_c - \phi| < \varepsilon$$

Proof Sketch: It suffices to consider $\mathbb{1}_E = \mathbb{1}_{\bigcup_{j=1}^N R_j}$.

By approximation of sets and smoothness of product of 1D mollifiers, we can approximate $\mathbb{1}_E$ by a smooth function. Further, the support of these functions are disjoint.

3.9.2 Dyadic Cubes

Let Q_k the dyadic cube (an n-D closed rectangle with same sides) of length $2^{-k}\mathbb{Z}$. Conveniently, $Q_{k+1} \subseteq Q_k$.

Let $E \subseteq \mathbb{R}^n$. Define

$$\begin{aligned} \underline{A}(E,k) &= \bigcup Q \subseteq Q_k & \qquad Q_k \subseteq E \\ \overline{A}(E,k) &= \bigcup Q \subseteq Q_k & \qquad Q_k \cap E \neq \emptyset \end{aligned}$$

clearly, $\underline{A} \subseteq \overline{A}$ so $m(\underline{A}(E,k)) \leq m(\overline{A}(E,k))$.

Hence, $\underline{A}(E,k) \nearrow_k$ and $\overline{A}(E,k) \searrow_k$.

Define $\underline{A}(E) = \lim_k \underline{A}(E,k)$ and $\overline{A}(E) = \lim_k \overline{A}(E,k)$. If $\underline{A}(E) = \overline{A}(E)$, then we say E has the same **Jordan content**

Lemma: Let U be open, then $U = \underline{A}(U)$. Moreover, U is a countable union of disjoint cubes

Proof: Fix $x \in U$. Define $\delta = \inf_{y \setminus U} |y - x| > 0$. We know $B_\delta(x) \subseteq U$. Therefore, there exists a cube containing x that is contained in $B_\delta(x)$.

For large k, $x \in \underline{A}(U_k) \subseteq \underline{A}(U)$.

The other direction is trivial.

By the nature of dyadic cubes, $\underline{A}(U)$ is a countable collection $\bigcup_{k=0}^\infty A(U,k) = \underline{A}(U)$.

We can rewrite

$$\bigcup_{k=0}^\infty A(U,k) = A(U,0) \bigcup_{k=1}^\infty \{\underline{A}(U,k) \setminus \underline{A}(U,k-1)\}$$

which ensures we have a disjoint union.

3.9.3 Change of Variables

Remark: one reasonably intuitive way is to:

- Show change of variables for Riemann integration

- Show that for smooth functions, Lebesgue and Riemann integration is the same.
- Show that for $\mathcal{L}^1$ functions, we can approximate by smooth functions.

where the main difficulty is generalizing the Riemann-Lebesgue correspondence past 1D.

However, we will take a more measure theoretic approach.

Theorem (Translation Invariance): Let $\tau_a(x) = x + a$ for $a \in \mathbb{R}^n$. Then,

1. $m(\tau_a(E)) = m(E)$
2. If $f : \mathbb{R}^n \to \mathbb{C}$ is either ≥ 0 or in $\mathcal{L}^1$,

$$\int f(x+a)\, d\mu = \int f\, d\mu$$

Proof sketch:

1) For rectangles, $m(A) = m(\tau_a(A))$. We claim this holds for general sets because τ_a is a bijection between $E(x+a)$ and E and also $A(x+a)$ and $A(x)$.

2) This gives us the result for the characteristic function which immediately gives us simple functions (and thus $\mathcal{L}^1$ functions by approximation).