MAT257: Real Analysis II

QiLin Xue

Fall 2021

Contents

1 Review: Inner Product, Cauchy–Schwarz, Triangle Inequality **2**

2 Review: Continuity, Distance, and Linear Algebra Review **4**

3 Intervals, Regions, Open and Closed Sets **7**

4 Compactness **11**

4.1 Extra Notes from Tutorial 13

5 Continuity **15**

5.1 Tutorial Notes 16

6 Differentiability **17**

1 Review: Inner Product, Cauchy–Schwarz, Triangle Inequality

Lectures 1 and 2 (Sept 10/13)

- Everything in the course will lead to Stokes' Theorem:

$$\int_C \mathrm{d}W = \int_{\partial C} W \tag{1}$$

This generalizes a well-known theorem in one-dimensional calculus, known as the Fundamental Theorem of Calculus:

$$\int_{[a,b]} F'(t) = F(b) - F(a) = \int_{\partial[a,b]} F \tag{2}$$

where $\partial[a,b] = \{\underbrace{b}_{+}, \underbrace{a}_{-}\}$.

- **Continuity in** $\mathbb{R}^n$: Recall that continuity in $\mathbb{R}$ is formally defined via $\delta - \epsilon$. However intuitively it means that if you wiggle the input by a tiny bit, you wiggle the output by a tiny bit.

A similar way can be used to view continuity in $\mathbb{R}^n$.

Definition: For $x, y \in \mathbb{R}^n$, **the standard (Euclidean) inner product** of x and y denoted

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i \tag{3}$$

The **norm squared** is defined as

$$|x|^2 = \langle x, x\rangle \tag{4}$$

and the **norm** of x is defined:

$$|x| = \sqrt{|x^2|} = \sqrt{\sum_{i=1}^{n} x_i^2} \tag{5}$$

Idea: There are multiple ways of defining $\mathbb{R}^n$. Some people will define it as the set of all column vectors while others define it as the set of all row vectors. In linear algebra, the distinction is important but in real analysis, this distinction is not too important.

- A **bilinear** function $f(x, y)$ means that the function is linear in each of the two variables. This means that

$$f(ax + by, z) = af(x, z) + bf(y, z) \tag{6}$$

and similarly the same thing for the other parameter.

- A **semi-linear** function $f(x)$ is one such that

$$f(ax) = |a|f(x) \tag{7}$$

Proposition 1: If $x, y, z \in \mathbb{R}^n$ and $a, b \in \mathbb{R}$, then:

0. $\langle\cdot,\cdot\rangle$ is bilinear and $|\cdot|$ is semi-linear. Also note that $\langle x, y\rangle = \langle y, x\rangle$.
1. $|x| \geq 0$ and $|x| = 0 \iff x = 0$.
2. Cauchy–Schwarz Inequality: $|\langle x, y\rangle| \leq |x||y|$ and equality holds if and only if x, y are dependent.
3. Triangle Inequality: $|x + y| \leq |x| + |y|$
4. Polarization Identity: $\langle x, y\rangle = \dfrac{|x+y|^2 - |x-y|^2}{4}$

Proof. We prove each part separately, and skip 0:

1. We have $|x| = \sqrt{\sum x_i^2} \geq 0$ since every x_i^2 is non-negative. Then $|x| = 0$ if and only if $x_i = 0$, i.e. $x = 0$.

2. Consider and note that $\left||y|^2x - \langle x, y\rangle y\right|^2 \geq 0$. This is equal to 0 if and only if the first term (a multiple of x) equals the second term (a multiple of y), which is equivalent to x, y being dependent.

   Next, note that

$$\begin{aligned} |s+t|^2 &= \langle s+t, s+t\rangle && (8)\\ &= \langle s, s\rangle + \langle s, t\rangle + \langle t, s\rangle + \langle t, t\rangle && (9)\\ &= |s|^2 + 2\langle s, t\rangle + |t|^2 && (10)\end{aligned}$$

   Using this result, we can simplify the earlier expression to get that

$$\begin{aligned} \left||y|^2x - \langle x, y\rangle y\right|^2 &= |y|^4|x|^2 + \langle x, y\rangle^2|y|^2 - 2|y|^2\langle x, y\rangle^2 && (11)\\ &= |y|^2(|y|^2|x|^2 - \langle x, y\rangle^2) && (12)\end{aligned}$$

   Since this quantity is non-negative, it follows that $|y|^2|x|^2 \geq \langle x, y\rangle^2$, which is what we wanted to show.

   Note that there is another part regarding equality, which will not be proven in these notes.

3. As both $|x+y|$ and $|x| + |y|$ are nonnegative, we can square both sides. It suffices to prove that

$$\begin{aligned} |x+y|^2 &\overset{?}{\leq} (|x| + |y|)^2 && (13)\\ \langle x+y, x+y\rangle &\overset{?}{\leq} |x|^2 + |y|^2 + 2|x||y| && (14)\\ \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle &\overset{?}{\leq} |x|^2 + |y|^2 + 2|x||y| && (15)\\ |x|^2 + 2\langle x, y\rangle + |y|^2 &\overset{?}{\leq} |x|^2 + |y|^2 + 2|x||y| && (16)\\ \langle x, y\rangle &\overset{?}{\leq} |x| \cdot |y| && (17)\end{aligned}$$

   which is true via Cauchy–Schwarz.

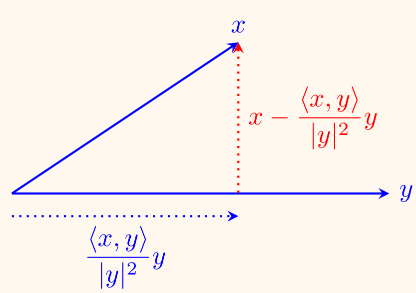
4. The proof is trivial. Expanding everything on the right hand side leads to the left hand side.

   Note: This property is important as it tells us that if we know how to compute norms, we can compute inner products.

□

Idea: In the proof for 2, a weird quantity $\left||y|^2x - \langle x, y\rangle y\right|^2$ was introduced. There is actually an intuition behind it. Recall that a geometric interpretation of Cauchy-Shwarz in $\mathbb{R}^2$ can be given as follows: $\langle x, y\rangle = |x||y|\cos\theta$ where θ is the "angle" between the two vectors x, y. This is smaller than $|x||y|$ since $\cos\theta \leq 1$.

Similarly, we want to find a generalized way to express something equivalent to $\cos\theta \leq 1$ in $\mathbb{R}^n$. One idea to show whether two vectors are dependent or not is to look at the separation between them, i.e. the length of the red line below.



Therefore, we have

$$\left|x - \frac{\langle x, y\rangle}{|y|^2}y\right| \geq 0 \tag{18}$$

Removing the fraction gives

$$\left||y|^2x - \langle x, y\rangle y\right| \geq 0. \tag{19}$$

2 Review: Continuity, Distance, and Linear Algebra Review

Lecture 2 and 3 (Sept 13 and 15) + Tutorial 1

- The properties in the previous section will be important when giving a formal definition of continuity.

Definition: If $x, y \in \mathbb{R}^n$, then:

$$d(x, y) = \text{“Distance between } x \text{ and } y\text{”} = |x - y| \tag{20}$$

Theorem:
1. d is **symmetric**, i.e. $d(x, y) = d(y, x)$.
2. d is **positive definite**, i.e. $d(x, y) \geq 0$ and $d(x, y) = 0 \iff x = y$.
3. Triangle Inequality: $d(x, z) \leq d(x, y) + d(y, z)$.

Proof. We prove each separately:

1. $d(x, y) = |x - y| = |-(y - x)| = |-1| \cdot |y - x| = |y - x| = d(y, x)$
2. $d(x, y) = 0 \iff |x - y| = 0 \iff x - y = 0 \iff x = y$
3. We need to check that

$$d(x, z) \overset{?}{\leq} d(x, y) + d(y, z) \tag{21}$$
$$|x - z| \overset{?}{\leq} |x - y| + |y - z| \tag{22}$$
$$|x - z| \overset{?}{\leq} |x - z| \tag{23}$$

where the third line comes from the previous triangle inequality. The last statement is true, and the steps are reversible, so we are done.

□

- This theorem is significant as these are the only properties that we need to know about distances to formally define continuity.

  Note: In a future section, we will use these properties to *define* a distance function (formally a metric), which is anything that satisfies these properties. This will allow us to generalize continuity to more abstract spaces.

- A note on notation. Our definition of a norm is known as the L^2 or *Euclidean norm,* i.e.

$$|x|_{L^2} = \sqrt{\sum x_i^2} \tag{24}$$

The L^1 norm can be defined as

$$|x|_{L^1} = \sum |x_i| \tag{25}$$

and the infinity norm:

$$|x|_\infty = \text{sup}|x_i|. \tag{26}$$

which is the maximum coordinate. Note that for a finite space, we can use max, but the use of supreme allows us to generalize to infinite spaces.

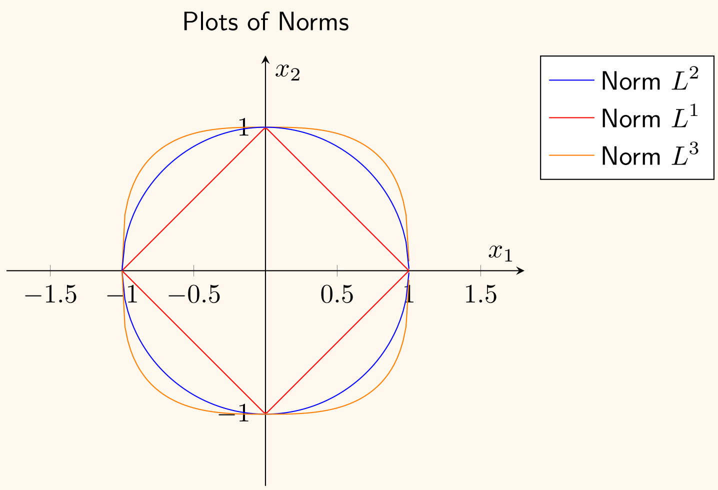
Idea: The p-norm can be defined as

$$|x|_{L^p} = \left(\sum x_i^p\right)^{1/p} \tag{27}$$

And it turns out that

$$|x|_\infty = \lim_{p \to \infty} |x|_{L^p}. \tag{28}$$

There is a nice geometric idea behind why this is the case. Consider $\mathbb{R}^2$. Then we can plot vectors with a norm of 1 in L^1, L^2, L^3 where the coordinate axes are x_1 and x_2.



We can see that as the dimension of the norm increases, the plot becomes closer and closer to a square, which can be represented by $\max\{x_1, x_2\}$.

Example 1: Suppose we have $\|\cdot\|_a$ on $\mathbb{R}^n$ and a norm $\|\cdot\|_b$ on $\mathbb{R}^n$. Suppose $\exists c > 0$ such that $\|x\|_b \leq c\|x\|_a$. Then if $f : \mathbb{R}^n \to \mathbb{R}$ is continuous with respect to the b-norm, then it is also continuous with respect to the a-norm.

Proof. We know that for any $\epsilon > 0$, we have $\|x - y\|_b < \delta_b \implies |f(x) - f(y)| < \epsilon$. Let us pick $\delta = \frac{\delta_b}{c}$. Then we have:

$$\|x - y\|_a < \frac{\delta_b}{c} \implies c\|x - y\|_a < \delta_b \tag{29}$$
$$\implies \|x - y\|_b \leq c\|x - y\|_a < \delta_b \tag{30}$$
$$\implies |f(x) - f(y)| < \epsilon \tag{31}$$

□

- Similarly, distances for L^1 and the infinity norms can be defined as $d_1(x, y) = |x - y|_1$ and $d_\infty(x, y) = |x - y|_\infty$.

  Exercise: Show that d_1 and d_∞ also satisfy the properties of a distance.

- There is a bijection from the linear map $\mathbb{R}^n \to \mathbb{R}^m$ and a $m \times n$ matrix:

$$\{T : \mathbb{R}^n \to \mathbb{R}^m\} \longleftrightarrow M_{m\times n}(\mathbb{R}) \tag{32}$$

  which is also a homomorphism. We can associate a matrix with any linear transformation, and any linear transformation is associated with a matrix. Here, the standard basis is used.

- Specifically, we have the map:

$$A \in M_{m\times n} \mapsto L_A(x) = Ax \tag{33}$$

  where $x \in \mathbb{R}^n$, and

$$T \mapsto M_T = \begin{pmatrix} Te_1 & Te_2 & \cdots & Te_n \end{pmatrix} \tag{34}$$

- We also need to show that this map is bijective, i.e both

$$L_{M_T} = T, \quad M_{L_A} = A \tag{35}$$

  are both satisfied.

- Furthermore, we can also show that this map is a homomorphism. Note that the set of linear transformations is itself a vector space. Both $A \mapsto L_A$ and $T \mapsto M_T$ is linear, so

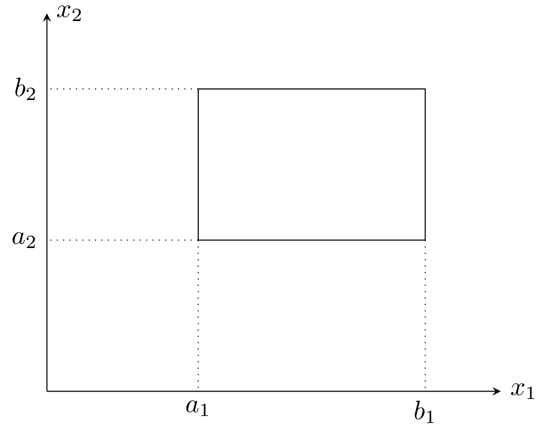
$$L_{aA+bB} = aL_A + bL_b \tag{36}$$
$$M_{aT+bS} = aM_T + bM_s. \tag{37}$$

- Furthermore, suppose we have two maps $T : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to \mathbb{R}^p$. To go from $\mathbb{R}^n$ to $\mathbb{R}^p$, we can take the composition
$$S \circ T \tag{38}$$
or
$$S \mathbin{/\!\!/} T \tag{39}$$
The bijection is a homomorphism, so we have
$$M_S M_T = M_{S \circ T} \tag{40}$$

3 Intervals, Regions, Open and Closed Sets

Lecture 3 and 4 (Sept 15, 20)

- In single-variable calculus, we typically focused on an interval $[a, b]$ on the real number line. Similarly, we can talk about intervals in $\mathbb{R}^2$ which can be represented as a rectangle:



- We can generalize to $\mathbb{R}^n$. Given $a_i \leq b_i$ for $i = 1, \dots, n$, we can define the **closed rectangle** corresponding to a_i, b_i:

$$R = \prod_{i=1}^{n} [a_i, b_i] = \{x \in \mathbb{R}^n : \forall i \, a_i \leq x_i \leq b_i\} \tag{41}$$

- Recall from set theory, if X and Y are sets, then $X \times Y = \{(x, y) : x \in X, y \in Y\}$ (also referred to as direct product in group theory). This is associative (up to isomorphism), so

$$(X \times Y) \times Z \neq X \times (Y \times Z) \tag{42}$$

  but

$$(X \times Y) \times Z \cong X \times (Y \times Z). \tag{43}$$

  As a result, while they are not equal strictly speaking, we can view them as the same.

- We can then view $\mathbb{R}^n$ as

$$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} = \{(x_1 \, \dots \, x_n) : x_i \in \mathbb{R}\} \tag{44}$$

- Products of intervals can be written as

$$\prod_{i=1}^{n} [a_i, b_i] = \{(x_1, \dots, x_n) : \forall i \, x_i \in [a_i, b_i]\} \tag{45}$$

- Likewise, there are also **open rectangles.** Specifically, the open rectangle defined by a_i, b_i is

$$\prod_{i=1}^{n} (a_i, b_i) = \{(x_1, \dots, x_n) \forall i \, x_i \in (a_i, b_i)\} \tag{46}$$

- There is a way to define continuity using open sets.

> **Definition**: The subset $A \subset \mathbb{R}^n$ is called "open" if:
>
> For every $a \in A$, there exists an open rectangle R, such that $a \in R \subset A$.

  Note that this isn't the only definition. Instead of surrounding a with a rectangle, we can surround a with an open ball B instead.

Definition: An open ball B is a ball of radius r centered at x_0:

$$B = B_r(x_0) = \{y \in \mathbb{R}^n : |x_0 - y| < r\} \tag{47}$$

This implicitly gives rise to the theorem:

Theorem: Defining "open" sets using rectangles is equivalent to defining "open" sets using balls.

Proof. The intuition of the proof is as follows:

1. Suppose A is open, so we can construct a rectangle R around $a \in A$.
2. In the rectangle, we can find a ball B such that $B \subset R$.

This shows one direction of the proof. The other direction is as follows:

1. Suppose A is open, so we can construct a ball B around $a \in A$.
2. In the ball, we can find a rectangle R such that $R \subseteq B$.

This is equivalent to showing that every open rectangle is open using the ball definition and that every open ball is open using the rectangle definition. □

- Note that balls depend on the norms that we use. Norms in L^1, L^2, L^∞ will all look different. We can use any of these norms to define a ball and still be valid in the theorem.
- We can also define closed sets.

Definition: B is **closed** if $\mathbb{R}^n \setminus B = B^C$ is open.

Theorem: We have the following theorems related to open and closed sets.
1. $\emptyset, \mathbb{R}^n$ are both closed and open.
2. Any union of open sets is open. Any intersection of closed sets is closed.
3. A finite intersection of open sets is open. A finite union of closed sets is closed.

Proof. We prove each separately:

1. To show that $\mathbb{R}^n$ is open, we just need to show that every point in $x \in \mathbb{R}^n$ is contained in a rectangle, i.e. $x \in \prod(x_i - 1, x_i + 1) \subset \mathbb{R}^n$. This also implies that $\emptyset$ is closed.

   To finish, we need to show that $\emptyset$ is open. Since there are no points in $\emptyset$, the condition holds[1]. This also implies that $\mathbb{R}^n$ is closed.

2. Suppose $\{A_\alpha\}_{\alpha \in I}$ is a collection of open sets, where I is an indexing set. We need to show

$$A = \bigcup_{\alpha \in I} A_\alpha = \{x : \exists \alpha \in I \text{ s.t. } x \in A_\alpha\} \tag{48}$$

   is open. Let $x \in A$. We need to find an open rectangle that surrounds x. Note that we can find an α such that $x \in A_\alpha$. Since A_α is open, there exists an open rectangle such that $x \in R \subset A_\alpha$.

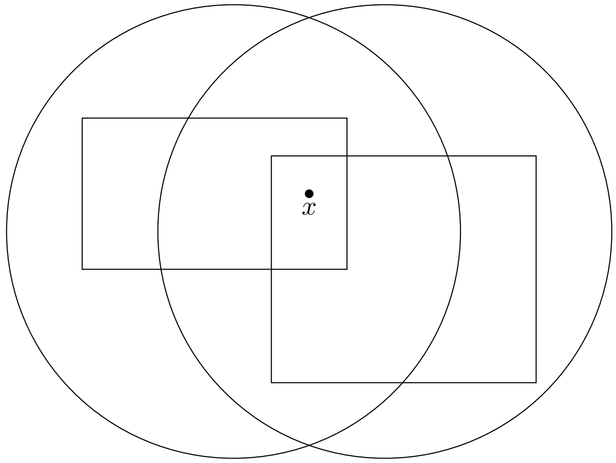
   Using De-Morgan's Laws, we can prove the second part. Suppose $\{B_\alpha\}_{\alpha \in I}$ is a collection of closed sets. We want to show that $\bigcap B_\alpha$ is closed. However,

$$\left(\bigcap B_\alpha\right)^C = \bigcup B_\alpha^C \tag{49}$$

   However, B_α^C is an open set, so by earlier, $\left(\bigcap B_\alpha\right)^C$ is open, so $\bigcap B_\alpha$ is closed.

3. We can use induction. We just need to show that the intersection of two open sets is open. We can visualize it as follows:

[1] This is analog to: Every horse in an empty set of horses has horns.



If we take the intersection of the two open rectangles, we see that we have a rectangle that surrounds the point x.

Lemma: The intersection of two open rectangles, if non-empty, is an open rectangle.

We can formalize as follows: Suppose $x \in A_1 \cap A_2$ by openness of A_i, there exists open rectangles R_i such that $x \in R_i \subset A_i$ for $i = 1, 2$. Then:

$$x \in \underbrace{R_1 \cap R_2}_{\text{open rectangle}} \subset A_1 \cap A_2. \tag{50}$$

To finish, we use induction. Suppose A_i, $i = 1, \ldots, n$ are open. Then:

$$\bigcap_{i=1}^{n} A_i = \left(\cap_{i=1}^{n-1} A_i\right) \bigcap A_n \tag{51}$$

By induction, the first part is open and by the base case, the intersection is open.

We can do something similar for the second part. Suppose B_i, $i = 1, \ldots, n$ is closed. Then:

$$\left(\bigcup_{i=1}\right)^{C} = \bigcap_{i=1}^{n} B_i^C \tag{52}$$

We've shown that the right hand side is open, so $\bigcup\limits_{i=1}$ must be closed.

□

- The following is important to relate intersections and unions.

Theorem: **De-Morgan's Laws:** If Y_α is any collection of subsets of some universe U. Then:

$$\left(\bigcup Y_\alpha\right)^C = \bigcap Y_\alpha^C \tag{53}$$

and

$$\left(\bigcap Y_\alpha\right)^C = \bigcup Y_\alpha^C \tag{54}$$

- Note that part (3) of the theorem on closed/open sets is not true if we take an infinite intersection. As a counterexample, consider

$$\bigcap_{n>0} \left(-\frac{1}{n}, 1 + \frac{1}{n}\right) = [0, 1] \tag{55}$$

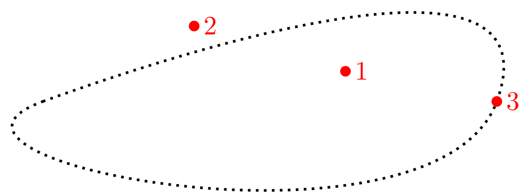
Again by De-Morgan, we can use this to construct an example of an infinite union of closed sets that is not closed.

- Given $A \subset \mathbb{R}^n$ and $x \in \mathbb{R}^n$, exactly one of the three things hold, i.e. a *trichotomy*:

1. $\exists$ open rectangle R such that $x \in R \subset A$.

2. $\exists$ open rectangle R such that $x \in R \subset A^C$
3. Every open rectangle R s.t. $x \in R$, has $R \cap A^C \neq \emptyset$ and $R \cap A \neq \emptyset$.

For example: 1,2,3 can be represented below:



Definition: Points that satisfy the trichotomy are known as:
1. Interior: $x \in \text{int}(A) = \mathring{A}$
2. Exterior: $x \in \text{ext}(A)$. The closure can be defined as $\text{cl}A = \bar{A} = (\text{ext}A)^C$.
3. Boundary: $x \in \text{Bd}(A)$.

- The closure of A can be defined as $\text{cl}(A) = \bar{A} = (\text{ext}A)^C$ and has the following property:

Theorem: We have $x \in \text{cl}(A)$ if and only if every open rectangle $R \ni x$ satisfies $R \cap A \neq \emptyset$.

- We also make the following claims:
  1. $\text{int}(A) \dot{\bigcup} \text{ext}(A) \dot{\bigcup} \text{Bd}(A) = \mathbb{R}^n$.
  2. $\text{cl}(A) = A \cup \text{Bd}(A)$
  3. $\text{int}(A) = A \setminus \text{Bd}(a)$

Example 2: Consider the set $[0, 1] \subset \mathbb{R}$. Then:
- $\text{int}(A) = (0, 1)$
- $\text{ext}(A) = (-\infty, 0) \cup (1, \infty)$
- $\text{cl}(A) = [0, 1]$
- $\text{Bd}(A) = \{0, 1\}$

4 Compactness

- We start with an open cover

Definition: An open cover of a set A is a collection $\{U_\alpha\}$ of open sets in $\mathbb{R}^n$ such that

$$\bigcup_{\alpha} U_\alpha \supset A \tag{56}$$

A subcover of $\{U_\alpha\}_{\alpha \in I}$ is a collection where α runs over $I' \subset I$ such that

$$\bigcup_{\alpha \subset I'} U_\alpha \supset A \tag{57}$$

Definition: A is called compact if every open cover of A has a finite sub-cover.

Example 3: If $F \subset \mathbb{R}^n$ is finite, then it is compact.

Example 4: $\mathbb{R}$ is not compact.

Proof. We just need to find a counterexample. Let $\mathbb{R} = \bigcup_{n \in \mathbb{Z}} (n-1, n+1)$.

Alternatively, we can write $\mathbb{R} = \bigcup_{n \in \mathbb{Z}} (-n, n)$. □

- We want to eventually classify all compact subsets of $\mathbb{R}^n$.

Theorem: The **Heine-Borel Theorem** tells us that $[0, 1]$ is compact.

Proof. Let $\{U_\alpha\}_{\alpha \in J}$ be an open cover of $I = [0, 1]$. To find a finite subcover, we will look for a point g such that we can find a finite subcover that covers $[0, g]$, and we want to push g as far to the right as possible.

Let us define $G = \{g \in [0, 1] : \exists J' \subset J \text{ s.t. } \bigcup_{\alpha \in J'} \supset [0, g]\}$ where J' is finite. We want to show that $1 \in G$.

Let γ be the furthest we can "push" g. Specifically, let $\gamma = \sup(G)$. This is well defined since G is clearly bounded $G \subset [0, 1]$ and is nonempty since $0 \in G$.

We claim that $\gamma = 1$. Suppose this is not true, i.e. $\gamma < 1$. If this is true, then there exists some open set U_β with $\beta \in J$ such that $\gamma \in U_\beta$. Since U_β is open, there exists g' and g'' such that $\gamma \in [g', g''] \subset U_\beta$. Therefore, $[0, g''] = [0, g'] \bigcup [g', g'']$. Since $[0, g']$ has a finite cover, and $[g', g'']$ is covered by U_β, then this has a finite cover, so $g'' \in G$. However, $\sup(G) \geq g'' > \gamma$, which leads to a contradiction.

□

Theorem: If $A \subset \mathbb{R}^n$ is compact and $B \subset \mathbb{R}^n$ is compact. Then:

$$A \times B \subset \mathbb{R}^{n+m} \tag{58}$$

is compact.

Proof. Suppose $\mathcal{U} = \{U_\alpha\}$ is an open cover of $A \times B$.

WLOG, each U_α is itself an open rectangle. The idea behind is that if we draw $A \times B$ as a rectangle, then every vertical line is a copy of B, which can be covered with finitely many covers. We then state and prove the following lemma.

Lemma 1: For every $x \in A$, we can find an open set $N_x \ni x$ such that $N_x \times B$ can be covered with finitely many of the U_α's.

Proof. Write $U_\alpha = V_\alpha \times W_\alpha$ where V_α and W_α are open rectangles in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively.

Consider $\{W_\alpha : x \in V_\alpha\}$ which covers B, which is compact, so there is a subcover $\{W_{\alpha_1}, \ldots, W_{\alpha_p}\}$ that covers B. Therefore:

$$U_{\alpha_1}, \ldots, U_{\alpha_p} \tag{59}$$

cover $\{x\} \times B$.

Let $N_x = \bigcap_{i=1}^{n} V_{\alpha_i} \subset V_{\alpha_i}$. We also have that for all α, $V_\alpha \ni x$. Now

$$\begin{aligned} N_x \times B &\subset \bigcup_{i=1}^{p} N_x \times W_{\alpha_i} && (60)\\ &\subset \bigcup_{i=1}^{p} V_{\alpha_i} \times W_{\alpha_i} && (61)\\ &= \bigcup_{i=1}^{p} U_{\alpha_i}, && (62)\end{aligned}$$

which is a finite open cover of $N_x \times B$. □

We know that $\{N_x\}_{x \in A}$ is an open cover of A. By compactness, we can find $x_1, \ldots, x_q$ such that

$$\bigcup_{j=1}^{q} N_{x_j} \supset A, \tag{63}$$

i.e.

$$\bigcup_{j=1}^{q} N_{x_j} \times B \supset A \times B \tag{64}$$

For each $j = 1, \ldots, q$, we can find U_{ji} where $i = 1, \ldots, p(j)$ such that

$$\bigcup_{i=1}^{p(j)} U_{ji} \supset N_{x_j} \times B. \tag{65}$$

Now,

$$\bigcup_{j=1}^{p} \bigcup_{i=1}^{p(j)} U_{ji} \supset A \times B. \tag{66}$$

□

Corollary 1: Any closed rectangles, i.e. regions in the form

$$R = \prod_{i=1}^{n} [a_i, b_i] \tag{67}$$

are compact.

Proposition 2: A closed subset of a compact set is compact.

Proof. Suppose C is compact and $B \subset C$ is closed.

This means that B^C is open. Suppose $\{U_\alpha\}$ is an open cover of B. Then

$$\{U_\alpha\} \cup \{B^C\} \tag{68}$$

is an open cover of C, so it has a finite subcover, which contains $U_{\alpha_1}, \ldots, U_{\alpha_p}$ and maybe B^C. If we just consider $U_{\alpha_1}, \ldots, U_{\alpha_p}$, and this is a finite cover of $B \subset C$. □

Corollary 2: Every closed and bounded subset of $\mathbb{R}^n$ is compact.

Definition: Bounded means that there exists $M \in \mathbb{R}$ such that for all $b \in B$, $|b| < M$. This is equivalent to saying that B is contained in some closed rectangle.

4.1 Extra Notes from Tutorial

- We can make the above corollary stronger (in $\mathbb{R}^n$):

Theorem: S is compact if and only if S is closed and bounded.

Proof. We have already shown that if S is closed and bounded, then S is compact. We now show the other direction.

- To show that S is bounded, note that by compactness there exists a finite subcover of S. Since these subcovers are finite, we can draw a ball or rectangle around this subcover, so it's bounded.
- To show that S is closed, we can show S^C is open. Let $x \in S^C$ and let $y \in S$ be the closest point to x contained in S. We can create an open ball centered at x with radius $\frac{1}{3}|x - y|$. This doesn't intersect S, so S^C is open and S is closed.

□

- This allows us to create three equivalent definitions of compactness. A set S is compact if and only if:
  - Every open cover has a finite subcover.
  - It is closed and bounded (only true for $\mathbb{R}^n$)
  - Every sequence of points in S has a convergent subsequence converging to a point S.
- Similarly, there are also three definitions of continuity.
  - We can define continuity using $\epsilon - \delta$.
  - A function is continuous if and only if pre-image of all open sets are open.
  - A function is continuous at a point p if whenever $\{a_n\}$ is a real sequence converging to p, the sequence $\{f(a_n)\}$ converges to $f(p)$.
- Finally, we can examine other ways of looking at the interior, boundary, exterior, and closure:
  - Int(S) is the largest open set in S
  - The closure, $\bar{S} = \text{cl}(S)$ is the smallest closed set containing S.

Example 5:

1. If C is closed:
   (a) Is $C \subseteq \text{cl}(\text{int}\, C)$? False, consider the empty set.
   (b) Is $C \supseteq \text{cl}(\text{int}\, C)$? True, the closure of a subset of S cannot contain points outside of S.
2. if U is open:
   (a) Is $U \subseteq \text{int}\,\text{cl}\, U$? True, since $\text{cl}\, U \subseteq \text{int}\,\text{cl}\, U$ and $U \subseteq \text{cl}\, U$.
   (b) Is $U \supseteq \text{int}\,\text{cl}\, U$? False, consider $(1, 2) \cup (2, 3)$.
3. (a) Is $\text{Bd}\, S \subseteq \text{Bd}\,\text{cl}\, S$? False, consider $(1, 2) \cup (2, 3)$.
   (b) Is $\text{Bd}\, S \supseteq \text{Bd}\,\text{cl}\, S$? True, $\text{Bd}\, S = \bar{S} \setminus \text{int}\, S$ and $\text{Bd}\, \bar{S} = \bar{S} \setminus \text{int}\, \bar{S}$. We have $\text{int}\, U \subseteq \text{int}\, \bar{U}$, so this is true.

- Given that f is a continuous function from $\mathbb{R}^n$ to $\mathbb{R}^m$, the following statements are true:
  1. f(open) $=$ open Consider $f(x) = c$
  2. f(closed) $=$ closed Consider $f(x) = \arctan(x)$
  3. f(bounded) $=$ bounded Consider $f(x) = \tan(x)$
  4. $f(\operatorname{int} S) = \operatorname{int}(f(S))$ Consider $f(x) = c$
  5. $f(\operatorname{Bd} S) = \operatorname{Bd} f(S)$ Consider $f(x) = c$

  but these two are true:
  1. $f(\operatorname{cl} S) = \operatorname{cl} f(S)$
  2. F(compact) $=$ compact

5 Continuity

- Review: Recall that if $C \subseteq \mathbb{R}^n$, then we can map C to its codomain via

$$F(C) = \{F(\gamma) : \gamma \in C\}. \tag{69}$$

Similarly, we can take the inverse image. For $D \in \mathbb{R}^m$:

$$F^{-1}(D) = \{\gamma \in \mathbb{R}^n : F(\gamma) \in D\} \tag{70}$$

Idea: It turns out that the inverse mapping is better behaved. For example,

$$F^{-1}(D_1 \cup D_2) = F^{-1}(D_1) \cup F^{-1}(D_2) \tag{71}$$

$$F^{-1}(D_1 \cap D_2) = F^{-1}(D) \cap F^{-1}(D_2) \tag{72}$$

$$F^{-1}(D^C) = F^{-1}(D)^C \tag{73}$$

However, only the following is true for the regular mapping:

$$F(C_1 \cup C_2) = F(C_1) \cup F(C_2) \tag{74}$$

- Consider the map $F : \mathbb{R}^n \to \mathbb{R}^m$, we have the following map

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \to \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} F_1(x_1 \cdots x_n) \\ \vdots \\ F_m(x_1 \cdots x_n) \end{pmatrix} \tag{75}$$

so we always get $F_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$, where F_i is known as a coordinate function of f. We can similarly write:

$$F_i = \pi_i \circ F = F \mathbin{/\!\!/} \pi_i \tag{76}$$

where π_i is the projection onto the i^{th} coordinate and $\circ$ denotes composition.

- Recall in one-dimensional calculus, the graph of a function $F : \mathbb{R} \to \mathbb{R}$ is given by $\Gamma_F = \{(x, f(x)) : x \in \mathbb{R}\} \subset \mathbb{R}^2$. We can generalize this to say that the graph of a function $F : \mathbb{R}^n \to \mathbb{R}^m$ lives in $\mathbb{R}^{n+m}$.
- We can also generalize limits.

Theorem: Suppose we have a function $F : \mathbb{R}^n \to \mathbb{R}^m$ and let $a \in A \subset \mathbb{R}^n$. Then

$$\lim_{x \to a} f(x) = b \tag{77}$$

means that $\forall \epsilon > 0, \exists \delta > 0$ s.t. $x \neq a \in B_\delta(a) \cap A \implies F(x) \in B_\epsilon(b)$. IF the limit exists, then it is unique.

- Note that we don't necessarily require $a \in A$. However in the case, then the limit may not be unique: in fact, the limit can be anything.
- As a result, it makes more sense to talk about the limit as x approaches $a \in \bar{A}$.

Definition: The function $F : A \subset \mathbb{R}^n \to \mathbb{R}^m$ is continuous at $a \in A$ if $\lim\limits_{x \to a} f(x) = f(a)$.

F is continuous on A if and only if it is continuous at every $a \in A$. This is equivalent to saying that for all a and $\epsilon > 0$, ther exists a $\delta > 0$ for all $x \in A$ such that $|x - a| < \delta \implies |f(x) - f(a)| < \epsilon$.

- There is another definition of continuity, which is equivalent to the above, but easier to work with.

Theorem: $F : \mathbb{R}^n \to \mathbb{R}^m$ is continuous if and only if for every open $V \subset \mathbb{R}^m$, the pre-image $F^{-1}(V)$ is also open.

Theorem: $F : A \subset \mathbb{R}^n \to \mathbb{R}^m$ is continuous if and only if for every open $V \subset \mathbb{R}^m$, there is an open $U \subset \mathbb{R}^n$ such that

$$F^{-1}(U) = U \cap A \tag{78}$$

- Aside: We can define that $B \subset A$ is **open in** A if there exists a U (that is open in $\mathbb{R}^n$) such that $B = U \cap A$. Note that being open in A has the same nice properties as open sets in general.

Theorem: A function $F : A \subset \mathbb{R}^n \to \mathbb{R}^m$ is continuous if and only if whenever $U \subset \mathbb{R}^m$ is open, there exists an open $V \in \mathbb{R}^n$ such that $F^{-1}(U) = V \cap A$.

Proof. We will only prove this theorem where $A = \mathbb{R}^n$, however the proof can easily extend to arbitrary A.

$\Longrightarrow$ Assume $U \subset \mathbb{R}^m$ is open, we need to show that $F^{-1}(U)$ is open. Pick $a \in F^{-1}(U)$, then $F(a) \in U$. We pick $\epsilon > 0$ such that $B_\epsilon(F(a)) \subset U$, by continuity, we can find $\delta > 0$ such that $F(B_\delta(a)) \subset B_\epsilon(F(a)) \subset U$, which is equivalent to saying $a \in B_\delta(a) \subset F^{-1}(U)$.

$\Longleftarrow$ Given $a \in \mathbb{R}^n$ and $\epsilon > 0$, consider $B_\epsilon(F(a))$ (which is open), so $F^{-1}(B_\epsilon(F(a))) \ni a$ is open. Since this is open, there exists $\delta > 0$ such that $a \in B_\delta(a) \subset F^{-1}(B_\epsilon(F(a)))$, so F is continuous. □

Theorem: Suppose we have a function $F : \mathbb{R}^n \to \mathbb{R}^m$ and $G : \mathbb{R}^m \to \mathbb{R}^p$. Given that F, G are continuous, then $G \circ F$ is continuous.

Proof. Given $U \subset \mathbb{R}^p$ is open, then

$$(G \circ F)^{-1}(U) = F^{-1}(G^{-1}(U)) = F^{-1}(\text{open}) = \text{open}. \tag{79}$$

□

Theorem: If $F : \mathbb{R}^n \to \mathbb{R}^m$ is continuous and $C \subset \mathbb{R}^n$ is compact, then $F(C)$ is compact.

Alternatively, a continuous image of a compact set is compact.

Proof. Given an open cover $\{U_\alpha\}$ of $f(C)$, then $\{F^{-1}(U_\alpha)\}$ is an open cover of C, hence it has a finite subcover. This finite subcover in $\mathbb{R}^n$ corresponds to a finite subcover in $\mathbb{R}^m$. □

Corollary 3: A continuous function on a compact set is bounded.

5.1 Tutorial Notes

Theorem: **Lebesque's Number Lemma:** If K is compact, then given an open cover $\mathcal{O}$, there exists $r > 0$ such that all balls of radius with less than r are contained in some $U \in \mathcal{U}$.

6 Differentiability

- We look at how we can define derivatives in multi-variable functions, specifically in $\mathbb{R}^n$.

Definition: $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a if and only if

$$f(a+h) = f(a) + Lh + o(h) \tag{80}$$

where $o(h) = \{e : \mathbb{R}^n \to \mathbb{R}^m : e(0) = 0, \lim\limits_{h\to 0} \dfrac{e(h)}{|h|} = 0\}$ and L is a linear transformation, which can be written as a matrix. We have abused some notation here.

- Some facts:
  - If f is constant, $f(x) = c$ for all x, then F is differentiable anywhere in its domain, and $f' = 0$.

    Proof. We want to write
    $$f(a+h) = f(a) + Lh + o(h). \tag{81}$$
    If we let $L = 0$ and $o(h) = 0$, then we are just left with $C = C$. ☐
  - If F is linear, say $F = L : \mathbb{R}^n \to \mathbb{R}^m$, then F is differentiable and $Df(a) = f$.

    Proof. Since f is linear, we have
    $$f(a+h) = f(a) + f(h) + 0 \tag{82}$$
    Therefore, $Df(a) = f$. ☐

    ]item Consider $s : \mathbb{R}^2 \to \mathbb{R}$ such that $s(x, y) = x + y$. We have that s is differentiable and $s' = s$.

    Note that the common name for this function is $+$, therefore $+' = + = \begin{pmatrix} 1 & 1 \end{pmatrix}$
- **Claim:** If $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are differentiable at a, then so is $f + g$ and $(f + g)' = f' + g'$.

Proof. Since f and g are differentiable,

$$f(a+h) = f(a) + f'(a)h + e_1(h) \tag{83}$$
$$g(a+h) = g(a) + g'(a)h + e_2(h) \tag{84}$$

Therefore,

$$\begin{aligned}(f+g)(a+h) &= f(a+h) + g(a+h) &(85)\\ &= f(a) + g(h) + (f'(a) + g'(a))h + e_1(h) + e_2(h) &(86)\\ &= (f+g)(h) + (f'(a) + g'(a))h + e_1(h) + e_2(h) &(87)\end{aligned}$$

Since $o(h)$ is a vector space, we have $e_1(h) + e_2(h) \in o(h)$. ☐

Theorem: Suppose $F : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a and $g : \mathbb{R}^m \to \mathbb{R}^p$ is differentiable at $\bar{a} = f(a)$, then $(g \circ f) : \mathbb{R}^n \to \mathbb{R}^p$ is differentiable at a and

$$D(g \circ f)(a) = (Dg)(f(a)) \cdot (Df)(a). \tag{88}$$

Alternatively,

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a) \tag{89}$$

where $\cdot$ denotes matrix multiplication. Note that we could restrict f to some open $A \ni a$ and g to some open $B \supset f(A)$.

Proof. We know that $f(a+h) = f(a) + f'(a)h + e_1(h)$ and $g(\bar{a} + \bar{h}) = g(\bar{a}) + g'(\bar{a})\bar{h} + e_2(\bar{h})$. Now,

$$\begin{aligned}(g \circ f)(a+h) &= g(f(a+h)) &(90)\\ &= g(\underbrace{f(a)}_{\bar{a}} + \underbrace{f'(a)h + e_1(h)}_{\bar{h}}) &(91)\\ &= g(\bar{a}) + g'(\bar{a})\bar{h} + e_2(\bar{h}) &(92)\end{aligned}$$

so the derivative is

$$g'(f(a))f'(a)h. \tag{93}$$

It is easy to check that $g'(\bar{a})e_1(h) + e_2(f'(a)h + e_1(h)) \in o(h)$. We can do this with a few lemmas. □

> **Lemma 2**: If A is a matrix and $e \in o(h)$ then $Ae \in o(h)$. Note that the two $o(h)$ are different since they may live in different dimensions.

> **Lemma 3**: If for small h, there exists some C such that $|\lambda(h)| < C|h|$ where $\lambda(h) = Lh + e(h)$. Then $e \circ \lambda$ is $o(h)$.

Showing these two lemmas are true will allow us to finish the proof.

Proof. Note that the first lemma is true since we can find constant C_1 such that $|Ah| \leq C_1|h|$. This is true for all h and this is proven in assignment 1.

First, we show that $(e \circ \lambda)(0) = 0$ and it remains to show that

$$\lim_{h\to 0} \frac{e(\lambda(h))}{|h|} = 0, \tag{94}$$

which is equivalent to saying that $\forall \epsilon > 0, \exists \delta > 0$ s.t. $|h| < \delta \implies |e(\lambda(h))| \leq \epsilon|h|$. To do this, suppose $|\lambda(h)| \leq C|h|$ on $B_{\delta_2}(0)$ and $|e(y)| \leq \frac{\epsilon}{C}|y|$ on $B_{\delta_1}(0)$. Set $\delta = \min\left\{\frac{\delta_1}{C}, \delta_2\right\}$ and get for $|h| < \delta$,

$$e(\lambda(h)) \leq \frac{e}{c}|\lambda(h)| \leq c|h| \tag{95}$$

□