MATH 247: Analysis 3

Instructor: Blake Madill

Spring 2021

Table of Contents

Part 2: Multivariable Calculus

Week 1 Notes.

Math 247: ① Introduction to Real Analysis
② Multivariable Calculus.

Module 1: Normed Vector spaces.

Definition 1.1.1 (normed vector space):
A normed vector space (NVS) is a vector space V (over $\mathbb{R}$) equipped with a function:

$$\|\cdot\|: V \to [0, \infty)$$

such that:

① $\|v\| = 0$ iff $v = 0$.

② For all $\alpha \in \mathbb{R}$, $v \in V$, $\|\alpha v\| = |\alpha| \|v\|$

③ (Triangle Inequality) For all $u, v \in V$,

$$\|u+v\| \leq \|u\| + \|v\|.$$

$\|\cdot\|$ is called a norm on V.

Remark 1.1.2: Geometric Motivations:
(notation) A NVS is represented by the pair $(V, \|\cdot\|)$.

① $\|v\|$: "the length of v" or "the distance between v and 0"

② $\|v-w\|$: "the distance between v and w".

Remark 1.1.3:
Real Analysis is the study of approximating objects releated to real numbers eg) $\mathbb{R}$, $\mathbb{R}^n$, sequences of functions from $\mathbb{R} \to \mathbb{R}$. Typically one approximates

bad, unruly elements (Eg $\mathbb{R}$) with nicely ordered elements (Eg $\mathbb{Q}$). The notion of distance in a NVS allows us to talk about approximation.

Example 1.1.4: $(\mathbb{R}, |\cdot|)$ (usual absolute value).

Example 1.1.5: $(\mathbb{R}, \|\cdot\|)$, where $\|\alpha\| = 3|\alpha|$.

Example 1.1.6: $V = \mathbb{R}^n$ (p norm).
For $p \geq 1$ and $v = (v_1, \dots v_n) \in \mathbb{R}^n$

$$\|v\|_p = \left(\sum_{i=1}^{n} |v_i|^p\right)^{1/p}$$

is a norm on $\mathbb{R}^n$.
When $p = 2$:

$$\|v\|_2 = \left(\sum_{i=1}^{n} v_i^2\right)^{1/2}$$

is called the Euclidean norm. This is the usual distance in $\mathbb{R}^n$.

Remark 1.1.7: We equip $\mathbb{R}^n$ with $\|\cdot\|_2$, unless stated otherwise.

Example 1.1.8: (infinity norm). $V = \mathbb{R}^n$

$$\|v\|_\infty = \max\{|v_1|, |v_2|, \dots, |v_n|\}.$$

Example 1.1.9: $V = \mathbb{R}^{\mathbb{N}} = \{(a_n)_{n=1}^{\infty} : a_i \in \mathbb{R}\}$.
$v = (v_1, v_2, \dots) \in V$.
For $p \geq 1$,

$$\|v\|_p = \left(\sum_{i=1}^{\infty} |v_i|^p\right)^{1/p}$$

$$\|v\|_\infty = \sup\{|v_1|, |v_2|, \dots\}.$$

Define:

① $\ell^p = \{v \in \mathbb{R}^{\mathbb{N}} : \|v\|_p < \infty\}$. ($\ell^p$ can be shown to be a subspace of V). $(\ell^p, \|\cdot\|_p)$ is a NVS.

② $\ell^\infty = \{v \in \mathbb{R}^{\mathbb{N}} : \|v\|_\infty < \infty\}$ (the set of all bounded sequences). $(\ell^\infty, \|\cdot\|_\infty)$ is a NVS.

Example 1.1.10:

$$V = C([a,b]) = \{f : [a,b] \to \mathbb{R} : f \text{ is continuous}\}.$$

① $p \geq 1$, $f \in V$.

$$\|f\|_p \equiv \left(\int_a^b |f(x)|^p dx \right)^{1/p}.$$

is a norm. This is called the integration norm.

② $\|f\|_\infty \equiv \sup\{|f(x)| : x \in [a,b]\}$

$$\overset{EVT}{=} \max\{|f(x)| : x \in [a,b]\}$$

is a norm. This is called the uniform norm.

Remark 1.1.11: Unless otherwise stated, we equip $C([a,b])$ with $\|\cdot\|_\infty$.

Module 2: Convergence.

Remark 1.2.1: V is a NVS. By a sequence is V we mean $(a_n)_{n=1}^{\infty} = (a_1, a_2, \dots)$, where each $a_i \in V$. We denote this by $(a_n) \subseteq V$.

Definition 1.2.2: V is a NVS. We say (a_n) converges to $a \in V$, written $a_n \to a$, if for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $\|a_n - a\| < \varepsilon$ for all $n \geq N$.

Definition 1.2.3: For $(a_n) \subseteq V$, if for all $a \in V$, $a_n \not\to a$, we say (a_n) diverges.

Example 1.2.4: $V = \ell^{\infty}$, $(a_n) \subseteq V$.

$$a_n = (1, \tfrac{1}{2}, \tfrac{1}{3}, \dots, \tfrac{1}{n}, 0, 0, \dots).$$

Claim: $a_n \to a$, where

$$a = (1, \tfrac{1}{2}, \tfrac{1}{3}, \dots).$$

Proof:

Let $\varepsilon > 0$ be given. Choose $N \in \mathbb{N}$ such that $1/N < \varepsilon$. For $n \geq N$, we have:

$$
\begin{aligned}
&\|a_n - a\|_\infty \\
&= \|(0, 0, \dots, -\tfrac{1}{n+1}, -\tfrac{1}{n+2}, \dots)\|_\infty \\
&= \sup\{0, |-\tfrac{1}{n+1}|, |-\tfrac{1}{n+2}|, \dots\} \\
&= \tfrac{1}{n+1} < \tfrac{1}{n} \leq \tfrac{1}{N} < \varepsilon.
\end{aligned}
$$

$\square$

Example 1.2.5: $V = l^\infty$, $(a_n) \subseteq V$ with

$$a_n = (\underbrace{1, 1, \cdots, 1}_{n \text{ times}}, 0, 0, \cdots).$$

Claim: $a_n \not\to a$ with

$$a = (1, 1, 1, \cdots).$$

Proof:

For any $n \in \mathbb{N}$,

$$\|a_n - a\|_\infty = 1. \qquad \square$$

Fact: (a_n) diverges

Definition 1.2.6: V is a NVS, $(a_n) \subseteq V$, $A \subseteq V$.

① We say A is bounded if there exists $M > 0$ such that $\|a\| \leq M$ for all $a \in A$.

② We say (a_n) is bounded if $\{a_1, a_2, \cdots\}$ is bounded. In other words, $\exists M > 0$, $\|a_n\| \leq M$ for all n.

Theorem 1.2.7 (convergence implies bounded): if $(a_n) \subseteq V$ is convergent, then (a_n) is bounded.

Proof:

Suppose $a_n \to a \in V$. So, there exists $N \in \mathbb{N}$ such that if $n \geq N$, then

$$\|a_n - a\| < 1.$$

For $n \geq N$, $\|a_n\| = \|(a_n - a) + a\| \leq \|a_n - a\| + \|a\|$

$$< 1 + \|a\|$$

Take $M = \max\{\|a_1\|, \cdots, \|a_{N-1}\|, 1 + \|a\|\}$.

We have that $\|a_n\| \leq M$ for all $n \in \mathbb{N}$ $\square$

Note: the converse is not true.

Proposition 1.2.8: for $(a_n), (b_n) \subseteq V$, $a_n \to a$, $b_n \to b$,

① $a_n + b_n \to a + b$

② if $\alpha \in \mathbb{R}$ then $\alpha a_n \to \alpha a$.

Module 3. Completeness.

Definition 1.3.1: We say $(a_n) \subseteq V$ is a Cauchy sequence if for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that:
$$\|a_n - a_m\| < \varepsilon$$
for all $n, m \geq N$.

Proposition 1.3.2: (convergence implies Cauchy). If $(a_n) \subseteq V$ is convergent then (a_n) is Cauchy.

Proof:

Let $\varepsilon > 0$ be given. There exists $N \in \mathbb{N}$ and $a \in V$ such that $\|a_n - a\| < \frac{\varepsilon}{2}$ for all $n \geq N$.

For $n, m \geq N$,
$$\begin{aligned}\|a_n - a_m\| &\leq \|a_n - a\| + \|a_m - a\| \\ &< \varepsilon.\end{aligned}$$
$\square$

Example 1.3.3 (being Cauchy does not implies convergence).

$c_{00} = \{(x_n) \in \ell^\infty : \exists N, \forall n \geq N, x_n = 0\}$.

$(c_{00}, \|\cdot\|_\infty)$ is a NVS

Let $a_n = (1, \frac{1}{2}, \dots, \frac{1}{n}, 0, 0, \dots)$.

$a = (1, \frac{1}{2}, \frac{1}{3}, \dots) \notin c_{00}$.

Then, $a_n \to a$ in ℓ^∞. $\Rightarrow (a_n) \subseteq \ell^\infty$ is Cauchy.

$\Rightarrow (a_n) \subseteq c_{00}$ is Cauchy.

Since $a_n \to a \notin c_{00}$ and limits are unique, we have $(a_n) \subseteq c_{00}$ diverges (see definition of divergence)

So we are motivated to study the NVS that Cauchy and convergence do agree.

Definition 1.3.4. Let V be a NVS.

① We say $A \subseteq V$ is complete if whenever $(a_n) \subseteq A$ is Cauchy then $\exists\, a \in A$ such that $a_n \to a$.

②. If V is complete, we call V a Banach space.

Example 1.3.5: $(c_{00}, \|\cdot\|_\infty)$ is not a Banach space.

Module 4. Banach Spaces.

Example 1.4.1: $(\mathbb{R}, |\cdot|)$ is a Banach space.

Remark 1.4.2:

$$v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n,\ 1 \le p < \infty.$$

① $\|v\|_p^p = |v_1|^p + \dots + |v_n|^p \le n \|v\|_\infty^p$

②. $\|v\|_\infty^p \le |v_1|^p + \dots + |v_n|^p = \|v\|_p^p$

③. $\|v\|_p \le \sqrt[p]{n}\, \|v\|_\infty$.

$\|v\|_\infty \lesssim \|v\|_p$.

④ So it suffices to show that $(\mathbb{R}^n, \|\cdot\|_\infty)$ is a Banach space to prove that $(\mathbb{R}^n, \|\cdot\|_p)$ are Banach spaces for all $1 \le p \le \infty$.

Proof:

Assume $(\mathbb{R}^n, \|\cdot\|_\infty)$ is a Banach space.

Let $1 \le p < \infty$ and let $(a_k) \subseteq \mathbb{R}^n$ be Cauchy with respect to $\|\cdot\|_p$. So, (a_k) is Cauchy with respect to $\|\cdot\|_\infty$ (Remark 1.4.2 ③) $\Rightarrow a_k \to a \in \mathbb{R}^n$ with respect to $\|\cdot\|_\infty$. $\Rightarrow a_k \to a$ with respect to $\|\cdot\|_p$ (Remark 1.4.2 ③) $\Rightarrow$ $(\mathbb{R}^n, \|\cdot\|_p)$ is a Banach Space.

Proposition 1.4.3: $(\mathbb{R}^n, \|\cdot\|_\infty)$ is a Banach Space.

Proof:

Suppose $(a_k) \subseteq \mathbb{R}^n$ is Cauchy. Say

$$a_k = (a_k^{(1)}, a_k^{(2)}, \dots, a_k^{(n)}).$$

for all $k \in \mathbb{N}$, where $a_k^{(i)} \in \mathbb{R}$.

Let $\varepsilon > 0$ be given. There exists $N \in \mathbb{N}$ such that $\|a_k - a_l\|_\infty < \varepsilon$ for all $k, l \geq N$. For $k, l \geq N$, and $1 \leq i \leq n$,

$$|a_k^{(i)} - a_l^{(i)}| \leq \|a_k - a_l\|_\infty < \varepsilon.$$

$\Rightarrow (a_k^{(i)})_{k=1}^\infty \subseteq \mathbb{R}$ is Cauchy for all $1 \leq i \leq n$.

Since $\mathbb{R}$ is complete.

$$a_k^{(i)} \to b_i \in \mathbb{R}$$

for all $1 \leq i \leq n$.

We claim $a_k \to (b_1, b_2, \dots, b_n)$.

Let $\varepsilon > 0$ be given. Fix $1 \leq i \leq n$. There exists $N_i \in \mathbb{N}$ such that $|a_k^{(i)} - b_i| < \varepsilon$ for all $k \geq N_i$.

Let $N = \max\{N_1, \dots, N_n\}$. For $k \geq N$,

$$\|a_k - (b_1, \dots, b_n)\|_\infty$$
$$= \max\{|a_k^{(i)} - b_i| : 1 \leq i \leq n\} < \varepsilon. \qquad \square$$

Proposition 1.4.4: ℓ^∞ is a Banach space.

Proof:

Let $(a_n) \subseteq \ell^\infty$ be Cauchy. For all $n \in \mathbb{N}$, we may write

$$a_n = (a_n^{(1)}, a_n^{(2)}, \dots)$$

with each $a_n^{(i)} \in \mathbb{R}$.

We claim For $i \in \mathbb{N}$, $(a_n^{(i)})$ is Cauchy.

Let $\varepsilon > 0$ be given.
There exisits $N \in \mathbb{N}$ such that

$$\| a_n - a_m \|_\infty < \varepsilon$$

for $n, m \geq N$.
Fix $i \in \mathbb{N}$. For $n, m \geq N$,

$$| a_n^{(i)} - a_m^{(i)} | \leq \sup \{ | a_n^{(i)} - a_m^{(i)} | : i \in \mathbb{N} \}.$$
$$= \| a_n - a_m \|_\infty < \varepsilon.$$

By completeness of $\mathbb{R}$,

$$a_n^{(i)} \to b_i . \quad (n \to \infty)$$

for all $i \in \mathbb{N}$.
Let $b = (b_1, b_2, \dots)$. we claim that $a_n \to b$.
Let $\varepsilon > 0$ be given. There exists $N \in \mathbb{N}$ such that

$$\| a_n - a_m \|_\infty < \frac{\varepsilon}{2}.$$

for all $n, m \geq N$.
We have that.

$$| a_n^{(i)} - a_m^{(i)} | < \frac{\varepsilon}{2}.$$

for all $n, m \geq N$ and all $i \in \mathbb{N}$.

Take $m \to \infty$, for $n \geq N$.

$$| a_n^{(i)} - b_i | \leq \frac{\varepsilon}{2}.$$

for all $i \in \mathbb{N}$.
So, $\| a_n - b \|_\infty \leq \frac{\varepsilon}{2} < \varepsilon$ for all $n \geq N$. $\square$

Week 2 Notes.

Module 1. Closed and Open Sets.

Remark 2.1.1: Big idea.

For a NVS, V, some subsets $A \subseteq V$ work nicer with limits/convergence than others. By understanding these "nice" subsets helps us understand convergence in V.

Definition 2.1.2: Let V be NVS.

1. A subset $C \subseteq V$ is closed whenever $(a_n) \subseteq C$, such that $a_n \to a \in V$, then $a \in C$.
2. We say $U \subseteq V$ is open if $V \setminus U$ is closed.
3. The collection

$$\{U \subseteq V : U \text{ is open}\}$$

is called the Topology on V. The study of open/closed sets on a space is called Topology.

Example 2.1.3: $\phi, V \subseteq V$ are open and closed.

Example 2.1.4: $[0,1) \subseteq \mathbb{R}$ is neither open nor closed.

Example 2.1.5: for $r>0$, $a \in V$.

The closed ball of radius r, centered at a,

$$\overline{B_r(a)} := \{x \in V : \|a-x\| \leq r\}.$$

is closed.

Proof:

Let $(a_n) \subseteq \overline{B_r(a)}$ such that $a_n \to b \in V$.

By definition:

$$\|a_n - a\| \leq r$$

for all $n \in \mathbb{N}$. since $a_n \to b$,

$$\|a_n - a\| \to \|b - a\|$$
$$\leq r \text{ (limit preserves order).}$$

$\Rightarrow b \in \overline{B_r(a)}$. Hence, $\overline{B_r(a)}$ is closed $\square$

Example 2.1.6 $\{x \in V : \|x - a\| \geq r\}$ is closed.

Example 2.1.7. The open ball.

$$B_r(a) := \{x \in V : \|x - a\| < r\}.$$

is open.

Example 2.1.8: $V = \ell^\infty$.

$$c_0 := \{(x_n) \in \ell^\infty : x_n \to 0\}$$

is closed.

Proof:

Let $(a_n) \subseteq c_0$ such that $a_n \to a \in \ell^\infty$.

Let

$$a_n = (a_n^{(1)}, a_n^{(2)}, \ldots),$$

for all $n \in \mathbb{N}$.

Hence, $\lim\limits_{k \to \infty} a_n^{(k)} = 0$ for all $n \in \mathbb{N}$.

Say $a = (b_1, b_2, \ldots)$.

Let $\varepsilon > 0$ be given. We may find $N_1, N_2 \in \mathbb{N}$.

① $\quad \|a_n - a\|_\infty < \frac{\varepsilon}{2}$ for $n \geq N_1$.

② $|a_{N_1}^{(k)}| < \frac{\varepsilon}{2}$ for $k \geq N_2$.

For $k \geq N_2$

$$|b_k| = |a_{N_1}^{(k)} - b_k - a_{N_1}^{k}| \leq |a_{N_1}^{(k)} - b_k| + |a_{N_1}^{(k)}|$$
$$\leq \|a_{N_1} - a\|_\infty + |a_{N_1}^{(k)}|$$
$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This shows $b_k \to 0$ and so $a = (b_1, b_2, \dots) \in c_0$.

Hence, c_0 is closed. □

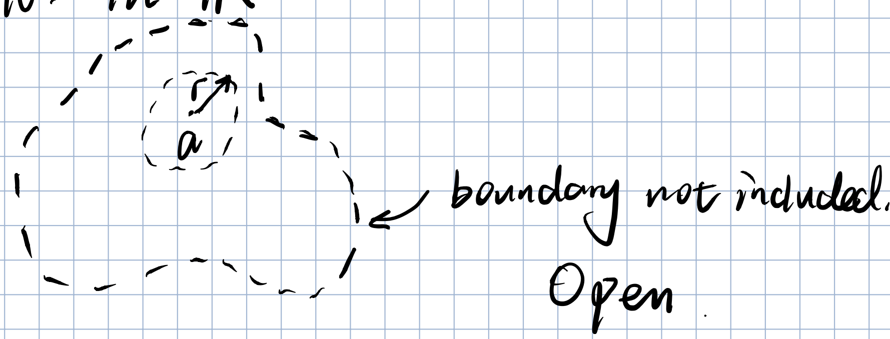
Proposition 2.1.9. Let V be NVS, $U \subseteq V$.

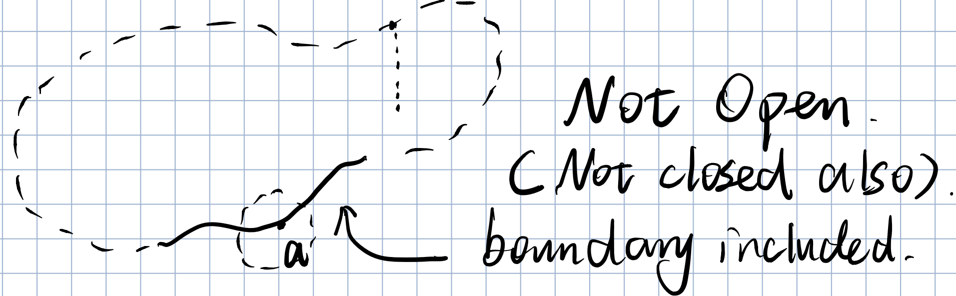
The following are equivalent:

① U is open.

② For all $a \in U$, there exists $r > 0$ such that
$$B_r(a) \subseteq U.$$

Example 2.1.10: In $\mathbb{R}^2$





Proof for proposition 2.1.9:

($\Rightarrow$) Assume U is open. Hence, $V \setminus U$ is closed.
For contradiction, assume $\exists\, a \in U$ such that $\nexists\, r > 0$ with $B_r(a) \subseteq U$.
For all $n \in \mathbb{N}$, we may find $a_n \in B_{\frac{1}{n}}(a)$ such that $a_n \notin U$.

Note:

1. $\|a_n - a\| < \frac{1}{n} \to 0 \Rightarrow a_n \to a$.
2. $(a_n) \subseteq V \setminus U$. $\leftarrow$ closed.

Therefore: $a \in V / U$, contradicting $a \in U$.

($\Leftarrow$). Assume for all $a \in U$, $\exists\, r > 0$ such that $B_r(a) \subseteq U$.

Claim: $V \setminus U$ is closed.
Let $(a_n) \subseteq V \setminus U$ such that $a_n \to a \in V$.
For contradiction, assume $a \in U$. We may find $r > 0$ such that $B_r(a) \subseteq U$. Since $a_n \to a$, we can find $N \in \mathbb{N}$ such that $\|a_N - a\| < r \Rightarrow$ $a_N \in B_r(a) \subseteq U \Rightarrow a_N \in U$, contradicting $a_N \in V \setminus U$.

Module 2: Closure and Interior.

Remark 2.2.1: Big Idea.

Given $A \subseteq V$, there are "naturally close" sets.

$U \subseteq A \subseteq C$ such that U is open and C is closed.

Proposition 2.2.2: Let V be NVS.

1. If $\{U_\alpha\}_{\alpha \in I}$ is a collection of open sets in V, then $U = \bigcup_{\alpha \in I} U_\alpha$ is open.

2. If $\{C_\alpha\}_{\alpha \in I}$ is a collection of closed sets in V. Then, $C = \bigcap_{\alpha \in I} C_\alpha$ is closed.

3. If $U_1, \ldots, U_n \subseteq V$ are open, then
$$\bigcap_{i=1}^{n} U_i$$
is open.

4. If $C_1, \ldots, C_n \subseteq V$ are closed, then
$$\bigcup_{i=1}^{n} C_i$$
is closed.

Proof Sketch.

1. Take $a \in \bigcup_{\alpha \in I} U_\alpha \Rightarrow \exists \alpha, a \in U_\alpha \Rightarrow \exists r > 0, B_r(a) \subseteq U_\alpha \Rightarrow \exists r > 0, B_r(a) \subseteq \bigcup_{\alpha \in I} U_\alpha$.

2. Consider $V \setminus \bigcap_{\alpha \in I} C_\alpha = \bigcup_{\alpha \in I} (\underbrace{V \setminus C_\alpha}_{\text{open}})$, $\rightarrow$ so this is open.
$\Rightarrow \bigcap_{\alpha \in I} C_\alpha$ is closed.

3. Take $a \in \bigcap_{i=1}^{n} U_i \Rightarrow \forall 1 \leq i \leq n, \exists r_i > 0$ such

that $B_{r_i}(a) \subseteq U_i$. $\underbrace{\text{Take } r = \min\{r_1, \dots, r_n\}}$
$\Rightarrow B_r(a) \subseteq \bigcap_{i=1}^{n} U_i$. So $\bigcap_{i=1}^{n} U_i$ is open.

This is where finite intersection comes from.

4. $V \setminus (\bigcup_{i=1}^{n} C_i) = \bigcap_{i=1}^{n} (\underbrace{V \setminus C_i}_{\text{open}}) \longrightarrow$ so this is open.

So, $\bigcup_{i=1}^{n} C_i$ is closed.

Example 2.2.3. $\bigcap_{n=1}^{\infty} \underbrace{(-\frac{1}{n}, \frac{1}{n})}_{\text{open}} = \underbrace{\{0\}}_{\text{not open}}$.

Example 2.2.4. $\bigcup_{n=1}^{\infty} \underbrace{[0, 1-\frac{1}{n}]}_{\text{closed}} = \underbrace{[0,1)}_{\text{not closed}}$.

Definition 2.2.5: Let $A \subseteq V$.

1. The closure of A is.

$$\bar{A} := \bigcap_{\substack{A \subseteq C \\ C \text{ closed}}} C$$ (this is closed as well).

2. The interior of A is:

$$\mathrm{Int}(A) := \bigcup_{\substack{U \subseteq A \\ U \text{ open}}} U$$ (this is open as well)

Remark 2.2.6:

1. $\bar{A}$ is the smallest closed set containing A.
2. $\mathrm{Int}(A)$ is the largest open set contained in A.

Definition 2.2.7: Let V be NVS.

1. If $A \subseteq V$, then a is called a limit point of A if $\exists (a_n) \subseteq A$ with $a_n \to a$.

2. If $A \subseteq V$, then an interior point of A is $a \in A$ such that $\exists r > 0$, $B_r(a) \subseteq A$.

Example 2.2.8: In $\mathbb{R}^2$.

a is limit point & interior point.
b is limit point but not interior point

Proposition 2.2.9: L $A \subseteq V$.

1. $\overline{A} = \{\text{limit points of } A\}$.
2. $\operatorname{Int}(A) = \{\text{Interior points of } A\}$.

Proof.

② Piazza.

①. Let $X = \{\text{limit points of } A\}$.

Claim: X is closed.

Let $(a_n) \subseteq X$ such that $a_n \to a \in V$.

For all $n \in \mathbb{N}$, $\exists\, b_n \in A$ such that $|a_n - b_n| < \frac{1}{n}$.

$\Rightarrow b_n = b_n - a_n + a_n \to 0 + a = a$.

$\Rightarrow a \in X$ since a is a limit point of A.

$\Rightarrow X$ is closed.

By definition, $\overline{A} \subseteq X$.

Let $x \in X$, so that $\exists (a_n) \subseteq A$ with $a_n \to x$.

Let $C \subseteq V$ be colsed such that $A \subseteq C$.
Then $(a_n) \subseteq C$ and so $x \in C$.
Hence $X \subseteq C$ for all $A \subseteq C$ with C closed.
$\Rightarrow X \subseteq \bar{A}$.
$\Rightarrow X = \bar{A}$. $\square$

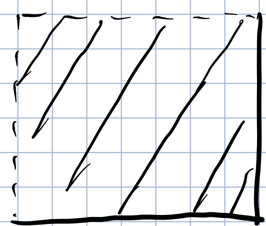
Remark 2.2.10. Summary.

① A closed $\Leftrightarrow$ A contains all its limit points.
$\Leftrightarrow A = \bar{A}$.

②. A open $\Leftrightarrow$ All points in A are interior points
$\Leftrightarrow A = \mathrm{Int}(A)$.

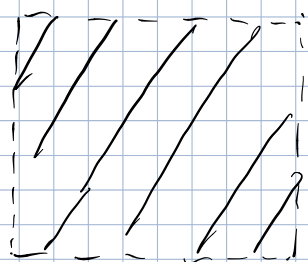
Module 3: Examples.

Example 2.3.1: In $\mathbb{R}^2$:

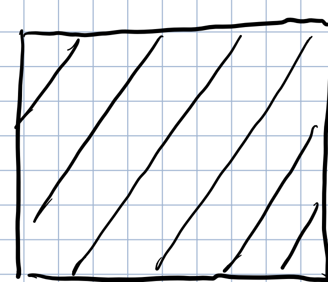


A.

Int(A):



$\overline{A}$:



Example 2.3.2: $A = \{(a_n) \in \ell^1 : a_n \in \mathbb{Q}\}$.

Claim: $\overline{A} = \ell^1$.

Remark 2.3.3: Let $x \in \ell^1$.

Assume $\forall \varepsilon > 0, \exists a \in A, \|x - a\|_1 < \varepsilon$.

Then, $x \in \overline{A}$.

$\forall n \in \mathbb{N}, \exists a_n \in A, \|x - a_n\|_1 < \frac{1}{n}$.

Then, $(a_n) \subseteq A$ and $a_n \to x \Rightarrow x \in \overline{A}$.

Proof for Example 2.3.2:

Let $x = (x_1, x_2, \dots) \in \ell^1$ and let $\varepsilon > 0$ be given.

By the density of rationals, $\forall n \in \mathbb{N}, \exists y_n \in \mathbb{Q}$ such that $|x_n - y_n| < \frac{\varepsilon}{2^n}$.

Consider $y = (y_1, y_2, \cdots)$

Then,

$$\|x-y\|_1 = \sum_{n=1}^{\infty} |x_n - y_n| < \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon.$$

Therefore, $x \in \overline{A}$, so $\ell^1 \subseteq \overline{A}$, $\ell^1 = \overline{A}$. $\square$

Example 2.3.4. $V = \ell^\infty$.

Claim: $\overline{c_{00}} = c_0$.

Proof:

We know that $c_{00} \subseteq c_0$ and c_0 is closed.

So, $\overline{c_{00}} \subseteq c_0$.

Let $x = (x_1, x_2, \cdots) \in c_0$ and let $\varepsilon > 0$ be given.

Sine $x_n \to 0$, $\exists N \in \mathbb{N}$ such that $|x_n| < \frac{\varepsilon}{2}$ for all $n \geq N$.

Let $y = (x_1, \ldots, x_{N-1}, 0, 0, \cdots) \in c_{00}$.

So, $\|x-y\|_\infty = \|(0, \cdots 0, x_N, x_{N+1}, \cdots)\|_\infty$

$$= \sup\{|x_k| : k \geq N\} \leq \frac{\varepsilon}{2} < \varepsilon.$$

So, $x \in \overline{c_{00}}$ and so $\overline{c_{00}} = c_0$.

(here we used Remark 2.3.3) $\square$

Module 4: More Properties.

Proposition 2.4.1: Let V be NVS.

1. $\overline{A\cup B} = \overline{A} \cup \overline{B}$

2. $\text{Int}(A\cap B) = \text{Int}(A) \cap \text{Int}(B)$.

3. $\overline{A\cap B} \subseteq \overline{A} \cap \overline{B}$.

4. $\text{Int}(A) \cup \text{Int}(B) \subseteq \text{Int}(A\cup B)$.

Proof:

① Since $\overline{A} \cup \overline{B}$ is closed and $A\cup B \subseteq \overline{A} \cup \overline{B}$,
$\overline{A\cup B} \subseteq \overline{A} \cup \overline{B}$.

Since $A, B \subseteq A\cup B$, we have $\overline{A}, \overline{B} \subseteq \overline{A\cup B}$,
so $\overline{A} \cup \overline{B} \subseteq \overline{A\cup B}$. $\square$

Example 2.4.2: $A = (0,1)$, $B = (1,2)$.

$\overline{A\cap B} = \overline{\phi} = \phi$.

$\overline{A} \cap \overline{B} = [0,1] \cap [1,2] = \{1\}$.

Example 2.4.3: $A = [0,1]$, $B = [1,2]$.

$\text{Int}(A\cup B) = (0,2)$

$\text{Int}(A) \cup \text{Int}(B) = (0,1) \cup (1,2)$

Proposition 2.4.4: Let $A \subseteq V$.

1. $\text{Int}(V\setminus A) = V \setminus \overline{A}$.

2. $\overline{V\setminus A} = V\setminus \mathrm{Int}(A)$.

Proof ①.

Since $V\setminus\overline{A} \subseteq V\setminus A$ and $V\setminus\overline{A}$ is open, we have $V\setminus\overline{A} \subseteq \mathrm{Int}(V\setminus A)$.

Observe that, $\underbrace{V\setminus\mathrm{Int}(V\setminus A)}_{\text{closed}} \supseteq V\setminus(V\setminus A) = A$.

and so $\overline{A} \subseteq V\setminus\mathrm{Int}(V\setminus A)$

$\Rightarrow V\setminus\overline{A} \supseteq \mathrm{Int}(V\setminus A)$.

Relpace A with $V\setminus A$ to prove ② $\square$

Definition 2.4.5. Let $A\subseteq V$.

The boundary of A is:

$$\partial(A) := \overline{A}\setminus\mathrm{Int}(A).$$

Proposition 2.4.6: Let $A\subseteq V$.

1. $\partial(A)$ is closed.
2. A is closed iff $\partial(A)\subseteq A$.

Proof Sketch:

1. $\partial(A) = \overline{A}\setminus\mathrm{Int}(A) = \underbrace{\overline{A}}_{\text{closed}} \cap (V\underbrace{\setminus\mathrm{Int}(A)}_{\text{closed}})$.

2. $(\Rightarrow)$ A is closed.

$\partial(A) \subseteq \overline{A} = A$.

$(\Leftarrow)$. $\partial(A)\subseteq A$

$\partial(A) = \overline{A}\setminus\mathrm{Int}(A)$

$$\Rightarrow \overline{A} = \underbrace{\partial(A)}_{\subseteq A} \cup \underbrace{\mathrm{Int}(A)}_{\subseteq A.}$$

$$\Rightarrow A \subseteq \overline{A} \subseteq A \Rightarrow A = \overline{A} \Rightarrow A \text{ is closed.}$$

Week 3 Notes.

Module 1. Compactness 1.

Remark 3.1.1: Big Idea.
Compact sets $A \subseteq V$ have nice "finiteness" topological properties. They may be infinite, but in some sence are "small".

Definition 3.1.2: Let V be NVS.
We say $C \subseteq V$ is compact if every $(a_n) \subseteq C$ has a subsequence $a_{n_k} \to a \in C$.

Example 3.1.3: $A \subseteq \mathbb{R}^n$ closed and bounded.
Claim: A is compact.
Let $(a_k) \subseteq A$.
Since A is bounded, (a_k) is bounded.
By assignment, $\exists (a_{k_l})$ such that $a_{k_l} \to a \in \mathbb{R}^n$.
Since A is closed, $a \in A$. $\square$

Example 3.1.4. $A = \{(1,0,0,\ldots),(0,1,0,0,\ldots), \ldots\} \subseteq \ell^\infty$
$(e_n) \subseteq A$ has no convergent sub sequence.
For $n \neq m$, $\|e_n - e_m\|_\infty = 1$.
So, A is not compact.

Example 3.1.5: $\overline{B_1(0)} \subseteq \ell^\infty$ closed, bounded.
But this is not compact. (see the sequence above)

Proposition 3.1.6: $C \subseteq V$ is compact Then C is closed and bounded.

Proof:

Let $C \subseteq V$ be compact.

① Claim: C is closed.

Let $(a_n) \subseteq C$ such that $a_n \to a \in V$.

$\exists (a_{n_k})$ such that $a_{n_k} \to b \in C$.

However, we must have $a = b \in C$ (subsequences of a convergent sequence converge to the same limit).

② Claim: C is bounded.

Suppose C is not bounded.

For all $n \in \mathbb{N}$ we may find $a_n \in C$ such that $\|a_n\| \geq n$. Consider $(a_n) \subseteq C$. Every subsequence of (a_n) is unbounded $\Rightarrow$ divergent, contradicting C is compact. $\square$

Theorem 3.1.7 [Heine-Borel Theorem].

A set $C \subseteq \mathbb{R}^n$ is compact iff it is closed and bounded (already proved).

Proposition 3.1.8: Let $C \subseteq V$ be compact.

If $A \subseteq C$ is closed, then A is compact.

Proof sketch:

$(a_n) \subseteq A \subseteq C \Rightarrow \exists (a_{n_k})$ converges in C.

Since A is closed, the limit belongs to A.

Module 2. Open Covers.

Remark 3.2.1: Goal:

Give an alternate description of compactness which exhibits the finiteness motivation from Module 1.

Definition 3.2.2: Let V be NVS, $A \subseteq V$.

① An open cover of A is a collection of open sets $\{U_\alpha : \alpha \in I\}$ such that

$$A \subseteq \bigcup_{\alpha \in I} U_\alpha .$$

It is called finite if $|I| < \infty$.

② A subset of an open cover of A, $\{U_\alpha : \alpha \in I\}$, which is also an open cover of A, is called a subcover of $\{U_\alpha : \alpha \in I\}$.

Example 3.2.3: $V = \mathbb{R}$, $A = [0,1]$

An open cover of A:

$$A \subseteq \bigcup_{\alpha \in [0,1] \cap \mathbb{Q}} (\alpha - \tfrac{1}{4}, \alpha + \tfrac{1}{4}).$$

A finite subcover:

$$A \subseteq (-\tfrac{1}{4}, \tfrac{1}{4}) \cup (0, \tfrac{1}{2}) \cup (\tfrac{1}{4}, \tfrac{3}{4}) \cup (\tfrac{1}{2}, 1) \cup (\tfrac{3}{4}, \tfrac{5}{4}).$$

Example 3.2.4: $V = \mathbb{R}^2$, $A = \mathbb{Z} \times \mathbb{Z}$.

$$A \subseteq \bigcup_{\alpha \in \mathbb{Z} \times \mathbb{Z}} B_{\frac{1}{2}}(\alpha).$$

There is no finite subcover.

Example 3.2.6: $V = \mathbb{R}$, $A = [0,1]$

$$A \subseteq \bigcup_{n \in \mathbb{N}} \left(\frac{1}{n}, 2\right).$$

There is no finite subcover.

Theorem 3.2.7: Let V be NVS, $A \subseteq V$.
Then, $A \subseteq V$ is compact iff every open cover of A has a finite subcover.
Will be proved in later modules.

Module 3. Compactness 2.

Lemma 3.3.1. Let V be NVS, $A \subseteq V$ be compact. Let $A \subseteq \bigcup_{\alpha \in I} U_\alpha$ be an open cover of A. There exists $R > 0$ such that for all $a \in A$, $B_R(a) \subseteq U_\alpha$ for some $\alpha \in I$.

Proof:

Suppose no such $R > 0$ exists. In particular, for all $n \in \mathbb{N}$, $\exists a_n \in A$ such that $B_{\frac{1}{n}}(a_n) \nsubseteq U_\alpha$ for all $\alpha \in I$.

Since $(a_n) \subseteq A$ and A is compact, $\exists a_{n_k} \to a \in A$. Say $a \in U_{\alpha_0}$, $\alpha_0 \in I$.

Pick $M \in \mathbb{N}$ such that $B_{\frac{2}{M}}(a) \subseteq U_{\alpha_0}$.

Moreover, since $a_{n_k} \to a$, we may find $N \in \mathbb{N}$ such that $a_{n_k} \in B_{\frac{1}{M}}(a)$ for $k \geq N$.

Then, for $k \geq N$, such that $n_k \geq M$, take

$$
\begin{aligned}
x \in B_{\frac{1}{M}}(a_{n_k}) \Rightarrow \|x-a\| &= \|x - a_{n_k} + a_{n_k} - a\| \\
&\leq \|x - a_{n_k}\| + \|a_{n_k} - a\| \\
&< \frac{1}{M} + \frac{1}{M} = \frac{2}{M} \\
&\Rightarrow x \in B_{\frac{2}{M}}(a).
\end{aligned}
$$

Therefore, $B_{\frac{1}{M}}(a_{n_k}) \subseteq B_{\frac{2}{M}}(a) \subseteq U_{\alpha_0}$.

Since $n_k > M$, $B_{\frac{1}{n_k}}(a_{n_k}) \subseteq B_{\frac{1}{M}}(a_{n_k}) \subseteq U_{\alpha_0}$.

Contradiction. $\square$

Remark 3.3.2: R is called the Lebesgue number.

Proposition 3.3.3 [Part 1] Let V be NVS.
If $A \subseteq V$ is compact, then every open cover of A has a finite subcover.

Proof:
Suppose $A \subseteq V$ is compact. Let $A \subseteq \bigcup_{\alpha \in I} U_\alpha$ be an open cover of A.

We may find $R > 0$ as in the Lemma.
If $\exists\, a_1, \cdots a_n$ such that $A \subseteq B_R(a_1) \cup \cdots \cup B_R(a_n)$
we are done.

So, suppose there is no such covering. Take:

$$a_1 \in A.$$
$$a_2 \in A,\ a_2 \notin B_R(a_1).$$
$$a_3 \in A,\ a_3 \in B_R(a_1) \cup B_R(a_2).$$
$$\vdots$$

Since $(a_n) \subseteq A$ and A is compact, (a_n) has a convergent subsequence. However, for $n < m$,

$$a_m \notin B_R(a_n).$$
$$\Rightarrow \|a_m - a_n\| \geq R.$$

$\Rightarrow (a_n)$ has no Cauchy subsequence.
$\Rightarrow (a_n)$ has no convergent subsequence.
Contradiction. $\square$

Module 4. Compactness 3.

Proposition 3.4.1: Let V be NVS. $A \subseteq V$. If every open cover of A has a finite subcover, then A is compact.

Lemma 3.4.2. Let V be NVS, $A \subseteq V$. Suppose every open cover of A has a finite subcover. If $A \subseteq \bigcup_{\alpha \in I} U_\alpha$, where each U_α is relatively open in A, then $\exists \alpha_1, \alpha_2, \dots, \alpha_n \in I$ such that $A \subseteq U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}$.

Proof.

Suppose $A \subseteq \bigcup_{\alpha \in I} U_\alpha$, where $U_\alpha = A \cap \mathcal{O}_\alpha$, ($\mathcal{O}_\alpha \subseteq V$ is open).

$\Rightarrow A \subseteq \bigcup_\alpha (A \cap \mathcal{O}_\alpha) = A \cap (\bigcup_\alpha \mathcal{O}_\alpha) \subseteq \bigcup_\alpha \mathcal{O}_\alpha$.

$\Rightarrow A \subseteq \mathcal{O}_{\alpha_1} \cup \cdots \cup \mathcal{O}_{\alpha_n}$ (by assumption).

$\Rightarrow A \subseteq U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}$. $\square$

Proof of Proposition 3.4.1:

Suppose $A \subseteq V$ such that every open cover of A has a finite subcover.

Consider $(a_n) \subseteq A$.

For $k \in \mathbb{N}$, consider $C_k = \overline{\{a_n : n \geq k\}} \cap A$.

We want to show $\bigcap_{k=1}^{\infty} C_k \neq \phi$.

Each C_k is relatively closed in A. Hence, every $U_k = A \setminus C_k$ is relatively open in A.

For contradiction, assume $\bigcap_{k=1}^{\infty} C_k = \phi$.

$A = A \setminus \phi = A \setminus (\bigcap C_k) = \bigcup (A \setminus C_k) = \bigcup_{k=1}^{\infty} U_k$.

By the lemma, there exists $\exists\, i_1 < i_2 < \cdots < i_l$ such that $A \subseteq U_{i_1} \cup U_{i_2} \cup \cdots \cup U_{i_l}$.

Since $C_1 \supseteq C_2 \supseteq \cdots$, we have $U_1 \subseteq U_2 \subseteq \cdots$.

So, $A \subseteq U_{i_l} \subseteq A$ (relatively open).

$\Rightarrow A = U_{i_l}$.

$\Rightarrow C_{i_l} = A \setminus U_{i_l} = \phi$,

Hower, $a_{i_l} \in C_{i_l} = \phi$, contradiction.

Thus, we may find $a \in \bigcap_{k=1}^{\infty} C_k$.

[Piazza]

So, we may find $n_1 < n_2 < n_3 < \cdots$ such that

$\|a_{n_k} - a\| < \frac{1}{k}$ for every $k \in \mathbb{N}$.

Hence, $(a_{n_k}) \subseteq A$ with $a_{n_k} \to a \in A$. $\square$

Week 4 Notes.

Module 1 Limits.

Remark 4.1.1: Goal:
Give a quick overview of limits of functions $f: A \to W$, $A \subseteq V$, where V and W are NVS.

Definition 4.1.2: $f: A \to W$, $A \subseteq V$, $a \in V$.
We say the limit of $f(x)$ as x approaches a is $w \in W$ if
① $a \in \overline{A \setminus \{a\}}$ and
② $\forall \varepsilon > 0$, $\exists \delta > 0$ such that if $x \in A$ with $0 < \|x - a\| < \delta$, then $\|f(x) - w\| < \varepsilon$.

We write $\lim\limits_{x \to a} f(x) = w$.

Note: w is unique.

Remark 4.1.3: The condition $a \in \overline{A \setminus \{a\}}$.
If $a \notin \overline{A \setminus \{a\}}$, then we call $a \in V$ an isolated point with respect to A.

If $a \notin \overline{A \setminus \{a\}}$, then $\exists\, r > 0$ such that $B_r(a) \cap A = \{a\}$ or $\emptyset$.
Proof sketch: suppose no such r exist, in particular for every $\frac{1}{n}$, $\exists\, a_n \in A \setminus \{a\}$ so that $a_n \in B_{\frac{1}{n}}(a)$.
$a_n \to a$, contradicting $a \notin \overline{A \setminus \{a\}}$.

Thus, there is no $x \in A$ with $0 < \|x - a\| < r$.

Proposition 4.1.4: $A \subseteq V$, $f, g, h: A \to \mathbb{R}$, $a \in \overline{A \setminus \{a\}}$.

① [limits preserve order].

If $\lim_{x \to a} f(x)$, $\lim_{x \to a} g(x)$ exist and $f(x) \leq g(x)$ for all $x \in A$, then $\lim_{x \to a} f(x) \leq \lim_{x \to a} g(x)$.

② [Squeeze Theorem].

If $f(x) \leq g(x) \leq h(x)$ and $\lim_{x \to a} f(x) = \lim_{x \to a} h(x) = L$ exist, then $\lim_{x \to a} g(x) = L$.

Remark 4.1.5:

We will use all limit laws freely. The proofs are nearly identical to $f: A \to \mathbb{R}$, $A \subseteq \mathbb{R}$ case.

Example 4.1.6:

$$\lim_{(x,y,z) \to (0,0,0)} \frac{xy^2 + x^2 z + xyz}{\sqrt{x^2 + y^2 + z^2}}$$

Solution: If $x \neq 0$.

$$0 \leq \left| \frac{xy^2 + x^2 z + xyz}{\sqrt{x^2 + y^2 + z^2}} \right| \leq \frac{|xy^2 + x^2 z + xyz|}{\sqrt{x^2}}$$

$$= \frac{|xy^2 + x^2 z + xyz|}{|x|} \leq \frac{|x|y^2 + x^2|z| + |x||y||z|}{|x|}$$

$$= y^2 + |x||z| + |y||z|.$$

Note: If $x = 0$, then $f(x,y,z) = 0$.

Since $\lim_{(x,y,z) \to (0,0,0)} y^2 + |x||z| + |y||z| = 0$.

We know the original limit is 0 by the squeeze theorem.

Example 4.1.7

$$\lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2+y^4}.$$

As $(\frac{1}{n},0)\to(0,0)$, $f(\frac{1}{n},0)=0\to 0$.

As $(\frac{1}{n^2},\frac{1}{n})\to(0,0)$, $f(\frac{1}{n^2},\frac{1}{n})=\frac{1}{n^4}/(\frac{1}{n^4}+\frac{1}{n^4})=\frac{1}{2}\to\frac{1}{2}$.

This limit does not exist.

Module 2 Continuity.

Remark 4.2.1: Notations.

Unless stated otherwise, we only consider $f: A \to W$, $A \subseteq V$ where V, W are NVS.

Definition 4.2.2:

We say that f is continuous at $a \in A$ if for all $\varepsilon > 0$, there exists $\delta > 0$ such that if $x \in A$ with $\|x - a\| < \delta$, then $\|f(x) - f(a)\| < \varepsilon$.

Remark 4.2.3:

① If $a \in \overline{A \setminus \{a\}}$, then f is continuous at a iff $\lim_{x \to a} f(x) = f(a)$

② If $a \notin \overline{A \setminus \{a\}}$, then f is continuous at a.

Proof Sketch: $\exists\, r > 0$, $B_r(a) \cap A = \{a\}$. Let $\varepsilon > 0$ be given. Choose $\delta = r$. If $x \in A$ and $\|x - a\| < \delta$, then $x = a$. So, $\|f(x) - f(a)\| = \|f(a) - f(a)\| = 0 < \varepsilon$.

Definition 4.2.4:

If f is continuous at all $a \in A$, then we say f is continuous.

Proposition 4.2.5: The following are equivalent:

① f is continuous.

② f preserves convergence.

③ $\forall$ open $U \subseteq W$ such that $f^{-1}(U)$ is relatively open in A.

Proof:

② ⟺ ③ from assignment.

① ⟺ ②

Suppose f is continuous and let $(a_n) \subseteq A$ such that $a_n \to a \in A$. Let $\varepsilon > 0$ be given.

There exists $\delta > 0$ such that if $x \in A$ and $\|x - a\| < \delta$, then $\|f(x) - f(a)\| < \varepsilon$.

Take $N \in \mathbb{N}$ such that $\|a_n - a\| < \delta$ for all $n \geq N$.

Then for $n \geq N$,

$$\|f(a_n) - f(a)\| < \varepsilon.$$

which shows $f(a_n) \to f(a)$.

Assume f preserves convergence.

For contradiction, suppose f is discontinuous at a.

There exists $\varepsilon > 0$ and $(a_n) \subseteq A$ such that $\|a_n - a\| < \frac{1}{n}$ but $\|f(a_n) - f(a)\| \geq \varepsilon$. Then $a_n \to a$ but $f(a_n) \not\to f(a)$.

Contradiction. $\square$

Example. 4.2.6: $P_i : \mathbb{R}^n \to \mathbb{R}$, $1 \leq i \leq n$.

$P_i(x_1, x_2, \dots, x_n) = x_i$ is continuous.

Proof:

Let $(a_k) \subseteq \mathbb{R}^n$ such that $a_k \to a \in \mathbb{R}^n$.

Say:

$$a_k = (a_k^{(1)}, a_k^{(2)}, \dots, a_k^{(n)}).$$
$$a = (b_1, b_2, \dots, b_n).$$

We know (by assignment), $a_k^{(i)} \to b_i$.

$\Rightarrow P_i(a_k) \to P_i(a)$.

So, P_i is continuous.

Proposition 4.2.7:

① If $f+g: A \to W$ are continuous, then $f+g$ and αf $(\alpha \in \mathbb{R})$ are continuous).

② If $f: A \to W_1$ is continuous and $g: B \to W_2$, $B \subseteq W_1$, is continuous, then $g \circ f$ is continuous.

Proof sketch:

① $(a_n) \subseteq A$ such that $a_n \to a$. Since f, g are continuous,

$$f(a_n) \to f(a).$$
$$g(a_n) \to g(a).$$

By the limit laws:

$$f(a_n) + g(a_n) \to f(a) + g(a).$$
$$\alpha f(a_n) \to \alpha f(a).$$

So, $f+g$, αf are continuous.

②. Let $(a_n) \subseteq A$ such that $a_n \to a$.

Since f is continuous, $f(a_n) \to f(a)$.

Since G is continuous,

$$g(f(a_n)) \to g(f(a)).$$

So, $f \circ g$ is continuous.

Module 3: Uniform Continuity.

Definition 4.3.1: $A \subseteq V$, $f: A \to W$.
We say f is uniformly continuous if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that if $x, a \in A$ with $\|x-a\| < \delta$, then $\|f(x) - f(a)\| < \varepsilon$.

Remark 4.3.2: Big idea.
The same δ works uniformly for all $a \in A$.
Note: uniformly continuous $\to$ continuous.

Example 4.3.3.
$f: A \to W$ is Lipschitz if $\exists M > 0$ such that
$$\|f(a) - f(b)\| \leq M \|a - b\|$$
for all $a, b \in A$.
Claim: Lipschitz $\Rightarrow$ uniformly continuous.
Proof:
Let $\varepsilon > 0$ be given. Choose $\delta = \frac{\varepsilon}{M}$
If $a, b \in A$ with $\|a-b\| < \delta$, then $\|f(a) - f(b)\| \leq M\|a-b\| < M\delta = \varepsilon$.

$\square$

Example 4.3.4: $f [0, \infty) \to \mathbb{R}$.
$$f(x) = x^2.$$
Claim: f is not uniformly continuous
Proof:
Let $a_n = n + \frac{1}{n}$, $b_n = n$, $n \in \mathbb{N}$.

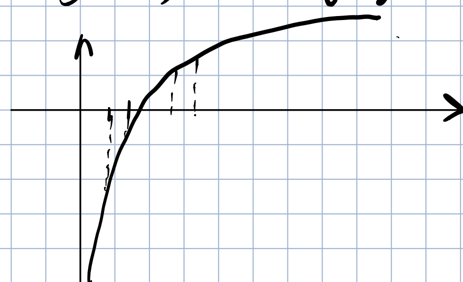
$$|a_n - b_n| = \frac{1}{n} \to 0.$$

$$|f(a_n) - f(b_n)| = \left(n + \frac{1}{n}\right)^2 - n^2 = 2 + \frac{1}{n^2} \geq 2.$$

$\square$

Example 4.3.5: $f: [0,1] \to \mathbb{R}$.

$f(x) = \ln(x)$



Proof:

Let $a_n = \frac{1}{n}$, $b_n = \frac{1}{n^2}$, $n \in \mathbb{N}$.

$$|a_n - b_n| = \frac{1}{n} - \frac{1}{n^2} = \frac{n-1}{n^2} \to 0.$$

$$|f(a_n) - f(b_n)| = \ln \frac{1}{n} - \ln \frac{1}{n^2} = \ln n \to \infty.$$

$\square$

Theorem 4.3.6:

If $C \subseteq V$ is compact and $f: C \to W$ is continuous, then f is uniformly continuous.

Proof Sketch:

Suppose f is not uniformly continuously.

$\exists \varepsilon > 0$, $(a_n), (b_n) \subseteq C$, $\|a_n - b_n\| < \frac{1}{n}$,

$\|f(a_n) - f(b_n)\| \geq \varepsilon$.

By compactness, $\exists\, a_{n_k} \to a \in C$.

$$\Rightarrow b_{n_k} = \underbrace{b_{n_k} - a_{n_k}}_{\to 0} + \underbrace{a_{n_k}}_{\to a} \to a.$$

By continuity: $f(a_{n_k}) \to f(a)$, $f(b_{n_k}) \to f(a)$.

$\Rightarrow \|f(a_{n_k}) - f(b_{n_k})\| \to 0$. Contradiction.

Module 4: Extreme Value Theorem.

Proposition 4.4.1:

If $C \subseteq V$ is compact, $f: C \to W$ is continuous, then $f(C)$ is compact.

Proof Sketch:

Take $(f(a_n)) \subseteq f(C)$, $a_n \in C$.

$(a_n) \subseteq C$. $\exists\ a_{n_k} \to a \in C$.

$\Rightarrow f(a_{n_k}) \to f(a) \in f(C)$.

Lemma 4.4.2.

If $A \subseteq \mathbb{R}$ is bounded and non-empty, then $\inf A, \sup A \in \overline{A}$.

Proof Sketch:

$\forall n \in \mathbb{N}$, $\exists\ \sup A - \frac{1}{n} < \overset{\in A}{a_n} \leq \sup A$

By squeeze theorem, $a_n \to \sup A$, so $\sup A \in \overline{A}$.

Theorem 4.4.3 [Extreme Value Theorem].

If $C \subseteq V$, $C \neq \phi$ and C is compact, $f: C \to \mathbb{R}$ is continuous, then there exists $a, b \in C$ such that

$$f(a) = \min f(C).$$
$$f(b) = \max f(C).$$

Proof:

$f(C)$ is compact by the proposition.

$\Rightarrow f(C)$ is closed and bounded.

$\sup f(C), \inf f(C) \in \overline{f(C)} = f(C)$.

$$f(a) = \inf f(C) = \min f(C).$$

$$f(b) = \sup f(c) = \max f(c). \qquad \square$$

Proposition 4.4.4:

Let $K \subseteq V$ be compact, W be NVS.

$C(K, W) = \{f : K \to W \text{ continuous}\}$ is a NVS.

When equipped with the uniform norm:

$$\|f\|_\infty = \max\{\|f(x)\| : x \in K\}.$$

Note: $f : K \to W$ is continuous

$\|\cdot\| : W \to \mathbb{R}$ is continuous.

$\Rightarrow \|f\| : K \to \mathbb{R}$ is continuous.

By the Extreme Value Theorem, $\max\{\|f(x)\| : x \in K\}$ exists.

Week 5 Notes

Module 1. Sequences of Functions.

Remark 5.1.1: Big Idea.
Given $A \subseteq V$ and W, and a sequence of functions $f_n: A \to W$, how can we make sense of (f_n) "converging" to some function $f: A \to W$?

Definition 5.1.2: $A \subseteq V$, $f_n: A \to W$, $f: A \to W$.
① We say f_n converges to f pointwise if $f_n(x) \to f(x)$ for all $x \in A$.
② We say f_n converges to f uniformly if for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $\|f_n(x) - f(x)\| < \varepsilon$ for all $n \geq N$ and $x \in A$.

Remark 5.1.3: Idea.
If $f_n \to f$ uniformly, then the same N works uniformly for all $x \in A$.

Remark 5.1.4: Notation.
Let $f_n, f: A \to W$, $A \subseteq V$.
$\|f_n - f\|_\infty := \sup\{\|f_n(x) - f(x)\| : x \in A\}$.
This is not necessarily a norm, because it could be ∞.
Note that:

$$f_n \to f \text{ uniformly} \iff$$

① $\|f_n - f\|_\infty < \infty$ eventually.
② $\|f_n - f\|_\infty \to 0$.

Example 5.1.5: $f_n: \mathbb{R} \to \mathbb{R}$.
$f_1(x) = x$, $f_n(x) = 0$ for all $n > 1$.
Then $f_n \to 0$ uniformly, even though $\|f_1 - 0\|_\infty = \infty$.

Module 2. Examples.

Given the following sequences of functions, find the pointwise limit and determine if the convergence is uniform.

Example 5.2.1: $f_n : (0,1) \to \mathbb{R}$.

$$f_n(x) = \frac{nx}{1+nx}.$$

For $x \in (0,1)$, $f_n(x) \to 1$. Therefore: $f_n \to 1$ pointwise.

Claim: this convergence is not uniform.

For $n > 1$.

$$\left| f_n\left(\frac{1}{n}\right) - 1 \right| = \frac{1}{2}.$$

So, $\|f_n - f\|_\infty \not\to 0$, and so the convergence is not uniform.

Example 5.2.2: $f_n : c_0 \to \mathbb{R}$, $f_n((a_k)) = a_n$

For $(a_k) \in c_0$, $f_n((a_k)) = a_n \to 0$ as $n \to \infty$.

So, $f_n \to 0$ pointwise

For $n \in \mathbb{N}$, $|f_n(\underbrace{1,1,1,\dots,1}_{n\text{-times}},0,0,0,\dots) - 0| = |1-0| = 1$.

$\Rightarrow \|f_n - 0\|_\infty \geq 1 \Rightarrow \|f_n - 0\|_\infty \not\to 0$, $f_n \not\to 0$ uniformly.

Example 5.2.3: $f_n : [0,1] \times [0,1] \to \mathbb{R}$.

$$f_n(a,b) = \frac{a^n}{n} + \frac{1}{b+n}$$

For $(a,b) \in [0,1] \times [0,1]$

$$f_n(a,b) = \underbrace{\frac{a^n}{n}}_{\to 0} + \underbrace{\frac{1}{b+n}}_{\to 0} \to 0.$$

So, $f_n \to 0$ pointwise.

Note:

$$|f_n(a,b) - 0| = \frac{a^n}{n} + \frac{1}{b+n} \leq \frac{1}{n} + \frac{1}{n} = \frac{2}{n}.$$

Therefore: $\|f_n - 0\|_\infty \leq \frac{2}{n} \to 0$, and so $f_n \to 0$ uniformly.

Module 3: Theorem A.

Example 5.3.1.

$$f_n : [0,1] \to \mathbb{R}.$$

$f_n(x) = x^n$ is continuous.

$$f_n \to \begin{cases} 1 & x=1 \\ 0 & x \neq 1 \end{cases} \quad \text{pointwise.}$$

Pointwise limit of continuous functions needs not to be continuous!

Theorem 5.3.2. (f_n), $f_n : A \to W$, $A \subseteq V$.
If f_n is continuous for all $n \in \mathbb{N}$ and $f_n \to f$ uniformly, then f is continuous.

Proof:
Let $(a_n) \subseteq A$ such that $a_n \to a$ and let $\varepsilon > 0$ be given.
We may find $N \in \mathbb{N}$ such that

$$\|f_N - f\|_\infty < \frac{\varepsilon}{3}$$

Since f_N is continuous, $\exists M \in \mathbb{N}$ such that
$|f_N(a_n) - f_N(a)| < \frac{\varepsilon}{3}$ for all $n \geq M$. Then for $n \geq M$,

$$\begin{aligned} |f(a_n) - f(a)| &\leq |f(a_n) - f_N(a_n)| + |f_N(a_n) - f_N(a)| + |f_N(a) - f(a)| \\ &\leq \|f - f_N\|_\infty + |f_N(a_n) - f_N(a)| + \|f_N - f\|_\infty \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Therefore, $f(a_n) \to f(a)$ and so f is continuous $\square$

Module 4.

Theorem 5.4.1. Let $A \subseteq V$ be compact, W be a Banach space, then $(C(A,W), \|\cdot\|_\infty)$ is a Banach space.

Proof:

Let $(f_n) \subseteq C(A,W)$ be Cauchy. Let $\varepsilon > 0$ be given. There exists $N \in \mathbb{N}$ such that

$$\|f_n - f_m\|_\infty < \varepsilon.$$

for all $n, m \geq N$.

For $x \in A$ and $n, m \geq N$,

$$\begin{aligned} &\|f_n(x) - f_m(x)\| \\ &\leq \|f_n - f_m\|_\infty < \varepsilon. \end{aligned} \qquad (*)$$

Hence, $(f_n(x)) \subseteq W$ is Cauchy. Since W is a Banach space, we know that

$$f_n(x) \to f(x) \in W$$

for some $f(x) \in W$.

We have constructed a function $f: A \to W$ such that $f_n \to f$ pointwise.

For $x \in A$ and $n \geq N$, $\lim_{m\to\infty} \|f_n(x) - f_m(x)\| \leq \varepsilon$.
(limits preserve order).

$\Rightarrow \|f_n(x) - f(x)\| \leq \varepsilon$
($f_m(x) \to f(x)$, $\|\cdot\|$ are continuous).

$\Rightarrow \|f_n - f\|_\infty \leq \varepsilon$.

So, $f_n \to f$ uniformly

By previous theorem, $f \in C(A, W)$ and so $f_n \to f$ in $C(A, W)$. This completes the proof. $\square$

Week 6 Notes.

Part 2: Multivariable Calculus.

Module 1. Partial Derivatives.

Remark 6.1.1. Idea.

Given a function $f: A \to \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$, develop a theory of multivariable differentiation using as much single variable as possible.

Definition 6.1.2.

A scalar function is a function of the form $f: A \to \mathbb{R}$, $A \subseteq \mathbb{R}^n$.

Remark 6.1.3:

If $f: A \to \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$ is a function, then $\exists f_1, f_2, \ldots f_m$ on A such that $f = (f_1, f_2, \ldots, f_m)$.

Example 6.1.4.

$f: \mathbb{R}^3 \to \mathbb{R}^2$.

$$f(x,y,z) = (\underbrace{xze^y}_{f_1(x,y,z)}, \underbrace{x^2+z^2}_{f_2(x,y,z)}).$$

Definition 6.1.5.

$A \subseteq \mathbb{R}^n$, $f: A \to \mathbb{R}$. Let $\{e_1, e_2, \ldots, e_n\}$ be the standard basis for $\mathbb{R}^n$. For $1 \le i \le n$, we define the i^{th} partial derivative of f at $a = (a_1, a_2, \ldots, a_n) \in A$ by:

$$f_{x_i}(a) = \frac{\partial f}{\partial x_i}(a)$$

$$:= \lim_{h \to 0} \frac{f(a + he_i) - f(a)}{h}$$

provided the limit exists.

Remark 6.1.6: Big picture.

$a \in A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$, we write $f(x_1, x_2, \dots, x_n)$.

① $f_{x_i}(a)$ is the derivative of f at a with respect to the variable x_i (treating x_j, $j \neq i$, as constants).

② $f_{x_i}(a)$ is the slope of the tangent line to the surface $y = f(x_1, x_2, \dots, x_n)$ which is parallel to e_i.

Example 6.1.7: $f : \mathbb{R}^2 \to \mathbb{R}$. $f(x, y)$.

$$a = (a_1, a_2) \in \mathbb{R}^2.$$

$$f_x(a) = \lim_{h \to 0} \frac{f(a + he_1) - f(a)}{h}$$

$$= \lim_{h \to 0} \frac{f(a_1 + h, a_2) - f(a_1, a_2)}{h}$$

Remark 6.1.8:

As usual, we can think of $\frac{\partial f}{\partial x_i}$ as a function

We write:

$$f_{x_i}(x_1, x_2, \dots, x_n)$$

or

$$\frac{\partial}{\partial x_i} f(x_1, x_2, \dots, x_n).$$

Example 6.1.9. $f : \mathbb{R}^3 \to \mathbb{R}$.

$$f(x, y, z) = xy^2 z + e^{xy}.$$

$$f_x(x,y,z) = y^2 z + y e^{xy}.$$

$$\frac{\partial f(x,y,z)}{\partial y} = 2xyz + x e^{xy}.$$

$$f_z(x,y,z) = xy^2$$

Definition 6.1.10.

$A \subseteq \mathbb{R}^n$, $f: A \to \mathbb{R}^m$, $f = (f_1, f_2, \ldots, f_m)$.

For $a \in A$

$$\frac{\partial f}{\partial x_i}(a) = f_{x_i}(a)$$

$$:= \left(\frac{\partial f_1}{\partial x_i}(a), \frac{\partial f_2}{\partial x_i}(a), \ldots, \frac{\partial f_m}{\partial x_i}(a)\right)$$

provided it exists.

Example 6.1.11: $f: \mathbb{R}^2 \to \mathbb{R}^3$.

$$f(x,y) = (2x^2 y,\ 4x,\ e^{xy}).$$

$$f_x(x,y) = (4xy,\ 4,\ y e^{xy})$$

$$f_y(x,y) = (2x^2,\ 0,\ x e^{xy}).$$

Module 2. Differentiability

Remark 6.2.1

Recall for single-variable functions.

$A \subseteq \mathbb{R}$, $f: A \to \mathbb{R}$. f is differentiable at $a \in A$.

iff.

① $a \in Int(A)$

② $\lim_{h \to 0} \frac{f(a+h)-f(a)}{h} = f'(a)$

exists

$\iff \lim_{h \to 0} \frac{f(a+h)-f(a)-mh}{h} = 0$.

for some $m \in \mathbb{R}$.

Remark 6.2.2.

Since $a \in Int(A)$, for small enough h, $a+h \in A$.

The limit above is well-defined.

Remark 6.2.3.

$T: \mathbb{R} \to \mathbb{R}$ is a linear transformation iff

$T(x) = mx$ for some $m \in \mathbb{R}$.

Remark 6.2.4: Idea

$$\lim_{h \to 0} \frac{f(a+h)-f(a)-\overset{T(h)}{mh}}{h} = 0.$$

f over $[a, a+h]$ (or $[a+h, a]$) can be approximated arbitrarily well by the line $T(x) = mx$.

Remark 6.2.5. Notation.

$\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) = \{T : \mathbb{R}^n \to \mathbb{R}^m : T \text{ is linear}\}$.

Definition 6.2.6:

$a \in A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}^m$, we say f is differentiable at $a \in A$ if

① $a \in \operatorname{Int}(A)$

② There exists $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ such that.

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - T(h)}{\|h\|} = 0.$$

Remark 6.2.7:

By ①, $f(a+h)$ is defined for small enough h.

Remark 6.2.8:

① Let $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Let B be the matrix of T relative to the standard ordered base. $T(x) = B(x)$ for all $x \in \mathbb{R}^n$.

② (assignment question).

Let $A \in M_{m \times n}(\mathbb{R})$.

$\|A\|_{op} = \sup\{\|Ax\| : x \in \mathbb{R}^n, \|x\| = 1\}$.

is a norm on $M_{m \times n}(\mathbb{R})$. We call it the operator norm.

③ If $A \in M_{m \times n}(\mathbb{R})$ and $x \in \mathbb{R}^n$. Then.

$$\|Ax\| \leq \|A\|_{op} \cdot \|x\|.$$

(already shown in assignment).

Theorem 6.2.9.

$a \in A \subseteq \mathbb{R}^n$, $f: A \to \mathbb{R}^m$. If f is differentiable at a, then f is continuous at a.

Proof Sketch:

f is differentiable at a

$\Rightarrow T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - T(h)}{\|h\|} = 0.$$

$$\Rightarrow \lim_{h \to 0} \frac{f(a+h) - f(a) - Bh}{\|h\|} = 0.$$

$\Rightarrow$ We may find $\delta > 0$ such that.

If $0 < \|h - 0\| = \|h\| < \delta$, then

$$\left\| \frac{f(a+h) - f(a) - Bh}{\|h\|} \right\| < 1.$$

$$\Rightarrow \|f(a+h) - f(a) - Bh\| < \|h\|.$$

$$\Rightarrow \|f(a+h) - f(a)\| - \|Bh\| < \|h\|.$$

$$\Rightarrow \|f(a+h) - f(a)\| < \|Bh\| + \|h\|.$$
$$\leq \|B\|_{op} \|h\| + \|h\|$$

As $h \to 0$, $\|B\|_{op} \|h\| + \|h\| \to 0$.

Therefore, $\lim_{h \to 0} f(a+h) = f(a)$.

Letting $x = a+h$. $\lim_{x \to a} f(x) = f(a)$, so f is continuous.

$\square$

Definition 6.2.10.
Let $U \subseteq \mathbb{R}^n$ be open, $f: U \to \mathbb{R}^m$.
We say that f is differentiable on U if f is differentiable at every point in U.

Module 3. Total Derivative.

Example 6.3.1. $f: \mathbb{R}^2 \to \mathbb{R}$.

$$f(x,y) = \begin{cases} (x^2+y^2)\sin\frac{1}{\sqrt{x^2+y^2}} & ((x,y) \neq (0,0)) \\ 0 & ((x,y) = (0,0)) \end{cases}$$

f is differentiable at $(0,0)$.

Consider $B = (0,0)$.

$$\lim_{h\to 0} \frac{f(0+h) - f(0) - Bh}{\|h\|} = \lim_{h\to 0} \frac{f(h)}{\|h\|}$$

$$= \lim_{(x,y)\to(0,0)} (x^2+y^2)\sin\frac{1}{\sqrt{x^2+y^2}} \cdot \frac{1}{\sqrt{x^2+y^2}}$$

$$= \lim_{(x,y)\to(0,0)} \sqrt{x^2+y^2}\sin\frac{1}{\sqrt{x^2+y^2}}$$

$= 0$ (Squeeze Theorem).

Remark 6.3.2: Problem: How to find the matrix B?

Remark 6.3.3: Investigation.

$a \in A \subseteq \mathbb{R}^n$, $f: A \to \mathbb{R}^m$

Suppose f is differentiable at a, so $\exists B \in M_{m\times n}(\mathbb{R})$ such that $\lim_{h\to 0} \frac{f(a+h) - f(a) - Bh}{\|h\|} = 0$.

Let $\{e_1, \cdots, e_n\}$ be the standard ordered basis.

As $t \to 0$, $te_i \to 0 \in \mathbb{R}^n$.

$$\Rightarrow \lim_{t\to 0} \frac{f(a+te_i)-f(a)-Bte_i}{|t|} = 0.$$

Therefore,

$$\lim_{t\to 0^+} \frac{f(a+te_i)-f(a)-tBe_i}{t} = 0.$$

and $\lim_{t\to 0^+} \frac{tBe_i}{t} = Be_i.$

$$\Rightarrow \lim_{t\to 0^+} \frac{f(a+te_i)-f(a)}{t} = Be_i.$$

$$\lim_{t\to 0^-} \frac{f(a+te_i)-f(a)-tBe_i}{-t} = 0.$$

$$\Rightarrow \lim_{t\to 0^-} \frac{f(a+te_i)-f(a)}{t} = Be_i.$$

$$\Rightarrow Be_i = \lim_{t\to 0} \frac{f(a+te_i)-f(a)}{t} = \frac{\partial f}{\partial x_i}(a).$$

$$= \left(\frac{\partial f_1}{\partial x_i}(a), \frac{\partial f_2}{\partial x_i}(a), \dots, \frac{\partial f_m}{\partial x_i}(a)\right).$$

$= (b_{1i}, b_{2i}, \dots, b_{mi})$. ($i$th column of B).

$\Rightarrow B = [b_{ij}]$ where $b_{ij} = \frac{\partial f_i}{\partial x_j}(a)$.

Definition 6.3.4:

$a \in A \subseteq \mathbb{R}^n$, $f: A \to \mathbb{R}^m$ We call the matrix

$$Df(a) = \left[\frac{\partial f_i}{\partial x_j}(a)\right]_{m\times n}.$$

the total derivative of f at a, provided it exists.

Theorem 6.3.5:

$a \in A \subseteq \mathbb{R}^n$, $f: A \to \mathbb{R}^m$. If f is differentiable at a, then:

① for all $1 \le j \le n$. $\frac{\partial f}{\partial x_j}(a)$ exists $(= Be_j)$.

② $\lim\limits_{h \to 0} \frac{f(a+h) - f(a) - Df(a)h}{\|h\|} = 0$.

Definition. 6.3.6: $a \in A \subseteq \mathbb{R}^n$, $f: A \to \mathbb{R}$.

We call $Df(a)$ the gradient of f at a and label it by $\nabla f(a)$.

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \cdots, \frac{\partial f}{\partial x_n}(a)\right).$$

Module 4. Continuous Partials.

Example 6.4.1.

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$$

$f_x(0,0) = 0$.
$f_y(0,0) = 0$ by symmetry.

$\nabla f(0,0) = 0$.
But f is not even continuous, so not differentiable.
$(\frac{1}{n}, \frac{1}{n}) \to 0$, $f(\frac{1}{n}, \frac{1}{n}) = \frac{1}{2} \to \frac{1}{2}$.

Theorem 6.4.2.
$U \subseteq \mathbb{R}^n$ open, $f: U \to \mathbb{R}$
If $a \in U$ and $\forall\, 1 \le j \le n$, if $\frac{\partial f}{\partial x_j}$ exist on U and are continuous at a, then f is differentiable at a.

Proof: Appendix.
Note: On assignment, this will be extended to $f: U \to \mathbb{R}^m$.

Example 6.4.3. $U = \mathbb{R}^2 \setminus \{(0,0)\}$.
Is $f: U \to \mathbb{R}$ given by $f(x,y) = \frac{\sin(xy)}{x^2+y^2}$ differentiable on U?

$$f_x(x,y) = \frac{(x^2+y^2)\cos(xy)y - 2x\sin(xy)}{(x^2+y^2)^2}.$$

$\hookrightarrow$ exists on U and is continuous on U.

By symmetry, $f_y(x,y)$ exists on U and is continuous.

By the Theorem, f is differentiable on U.

Remark 6.4.4.

Do not use the false converse of this theorem!

If $f: U \to \mathbb{R}$ is differentiable at $a \in U$, then there may exist a $\frac{\partial f}{\partial x_j}$ which is discontinuous at a.

Example 6.4.5: [Appendix].

$$f(x,y) = \begin{cases} (x^2+y^2)\sin\frac{1}{\sqrt{x^2+y^2}} & ((x,y)\neq(0,0)). \\ 0. & ((x,y)=(0,0)). \end{cases}$$

Week 6 Appendix

Theorem. Let $a \in U \subseteq \mathbb{R}^n$ be open and let $f : U \to \mathbb{R}$ be a function. If all of the partial derivatives of f exist on U (ie. the total derivative exists on U) **and are continuous at** a, then f is differentiable at a.

Proof. Suppose every partial derivative of f exists on U and that every partial derivative is continuous at a.
Suppose $a = (a_1, a_2, \ldots, a_n)$. Since U is open there exists $r > 0$ such that $B_r(a) \subseteq U$. For any $h = (h_1, h_2, \ldots, h_n) \neq 0$ such that $a + h \in B_r(a)$,

$$\begin{aligned} f(a+h) - f(a) &= f(a_1+h_1, \ldots, a_n+h_n) - f(a_1, \ldots, a_n) \\ &= f(a_1+h_1, \ldots, a_n+h_n) - f(a_1, a_2+h_2, \ldots, a_n+h_n) \\ &+ f(a_1, a_2+h_2, \ldots, a_n+h_n) - f(a_1, a_2, a_3+h_3, \ldots, a_n+h_n) \\ &+ f(a_1, a_2, a_3+h_3, \ldots, a_n+h_n) - f(a_1, a_2, a_3, a_4+h_4, \ldots, a_n+h_n) \\ &\vdots \\ &+ f(a_1, \ldots, a_{n-1}, a_n+h_n) - f(a_1, a_2, \ldots, a_n). \end{aligned}$$

However, by the singe variable Mean Value Theorem, for every $1 \leq j \leq n$ there exists c_j between a_j and $a_j + h_j$ such that

$$\frac{f(a_1, \ldots, a_{j-1}, a_j+h_j, \ldots, a_n+h_n) - f(a_1, \ldots, a_j, a_{j+1}+h_{j+1}, \ldots, a_n+h_n)}{a_j+h_j-a_j}$$

$$= \frac{\partial f}{\partial x_j}(a_1, \ldots, a_{j-1}, c_j, a_{j+1}+h_{j+1}, \ldots, a_n+h_n).$$

Putting all of this mess together,

$$f(a+h) - f(a) = \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a_1, \ldots, a_{j-1}, c_j, a_{j+1}+h_{j+1}, \ldots, a_n+h_n).$$

Now, for $1 \leq j \leq n$ let

$$\delta_j := \frac{\partial f}{\partial x_j}(a_1, \ldots, a_{j-1}, c_j, a_{j+1}+h_{j+1}, \ldots, a_n+h_n) - \frac{\partial f}{\partial x_j}(a_1, \ldots, a_n),$$

and $\delta = (\delta_1, \ldots, \delta_n)$. Then,

$$f(a+h) - f(a) - \nabla f(a) \cdot h = h \cdot \delta.$$

Since all of the partials are continuous at a, as $h \to 0$, each $\delta_j \to 0$, and so $\delta \to 0$ in $\mathbb{R}^n$. Therefore

$$0 \leq \lim_{h \to 0} \frac{|f(a+h) - f(a) - \nabla f(a) \cdot h|}{||h||} = \lim_{h \to 0} \frac{|\delta \cdot h|}{||h||} \leq \lim_{h \to 0} \frac{||\delta|| \cdot ||h||}{||h||} = 0.$$

Note that in the last inequality, we used the Cauchy-Schwarz inequality from linear algebra!. Therefore

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - \nabla f(a) \cdot h|}{||h||} = 0$$

and so

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - \nabla f(a) \cdot h}{||h||} = 0$$

as well. This exactly means that f is differentiable at a.

$\square$

Example. *Consider* $f : \mathbb{R}^2 \to \mathbb{R}$ *given by*

$$f(x,y) = \begin{cases} (x^2+y^2)\sin\left(\frac{1}{\sqrt{x^2+y^2}}\right) & \textit{if } (x,y) \neq (0,0) \\ 0 & \textit{if } (x,y) = (0,0). \end{cases}$$

From the module videos, we know that f is differentiable at $(0,0)$. Observe that

$$f_x(x,y) = 2x \sin\left(\frac{1}{\sqrt{x^2+y^2}}\right) - \cos\left(\frac{1}{\sqrt{x^2+y^2}}\right) \frac{x}{\sqrt{x^2+y^2}},$$

for all $(x,y) \neq (0,0)$. Now, we see that $(1/n, 0) \to (0,0)$ but

$$f_x(1/n, 0) = \frac{2}{n}\sin(n) - \cos(n)$$

diverges. Therefore f_x is not continuous at $(0,0)$.

Week 7 Notes.

Module 1 Differentiation Rules.

Remark 7.1.1: General Strategy.

$f: U \to \mathbb{R}^n$, $a \in U \subseteq \mathbb{R}^n$. U is open.

Suppose $\exists A \in M_{m\times n}(\mathbb{R})$ such that

$$\lim_{h\to 0} \frac{f(a+h)-f(a)-Ah}{\|h\|} = 0.$$

Then:

① f is differentiable at $a \in U$.

② $$\lim_{h\to 0} \frac{f(a+h)-f(a)-Ah}{\|h\|}$$

$$= \lim_{h\to 0} \frac{f(a+h)-f(a)-Df(a)h}{\|h\|}$$

$$\Rightarrow \lim_{h\to 0} \frac{(Df(a)-A)h}{\|h\|} = 0.$$

$$\Rightarrow \lim_{t\to 0^+} \frac{(Df(a)-A)te_i}{|t|} = 0.$$

$\Rightarrow (Df(a)-A)e_i = 0.$

$\Rightarrow Df(a)-A = 0.$

$\Rightarrow Df(a) = A$

Proposition 7.1.2: [Sum and Scalar Multiplication rule].

$f, g: A \to \mathbb{R}^m$, differentiable at $a \in A \subseteq \mathbb{R}^n$.

For $\alpha \in \mathbb{R}$, $f+\alpha g$ is differentiable at a and.

$$D(f+\alpha g)(a) = Df(a) + \alpha Dg(a).$$

Proof:

Let $P = f + \alpha g$

Then,

$$\lim_{h\to 0} \frac{P(a+h) - P(a) - (Df(a) + \alpha Dg(a))h}{\|h\|}$$

$$= \lim_{h\to 0} \frac{f(a+h) - f(a) - Df(a)h}{\|h\|} + \alpha \lim_{h\to 0} \frac{g(a+h) - g(a) - Dg(a)h}{\|h\|}$$

$$= 0 + \alpha 0 = 0.$$

Proposition 7.1.3. [Dot product rule].

$f, g : A \to \mathbb{R}^m$ differentiable at $a \in A \subseteq \mathbb{R}^n$.

Consider $f \cdot g : A \to \mathbb{R}$ defined by

$$(f \cdot g)(x) = f(x) \cdot g(x).$$

Then, $f \cdot g$ is differentiable at a and

$D(f \cdot g)(a) = g(a) Df(a) + f(a) Dg(a)$.

Proof: Appendix.

Module 2. Chain Rule.

Theorem 7.2.1.

$A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}^m$

$f: A \to \mathbb{R}^m$, $g: B \to \mathbb{R}^k$, $f(A) \subseteq B$.

If f is differentiable at $a \in A$, and g is differentiable at $f(a) \in B$, the $g \circ f$ is differentiable at a, with.

$$D(g \circ f)(a) = Dg(f(a)) \cdot Df(a).$$

Proof: Appendix.

Example 7.2.2. Let $f(x, y, z)$ be real-valued, differentiable. $x(t_1, t_2)$, $y(t_1, t_2)$, $z(t_1, t_2)$.

Then:

$$P(t_1, t_2) = (x(t_1, t_2), y(t_1, t_2), z(t_1, t_2)).$$

is differentiable.

Moreover:

$$D(f \circ p)(t_1, t_2) = Df(p(t_1, t_2)) Dp(t_1, t_2).$$

$$\Rightarrow \nabla f(t_1, t_2) = \nabla f(x, y, z) \cdot Dp(t_1, t_2)$$

$$\Rightarrow \left[\frac{\partial f}{\partial t_1}, \frac{\partial f}{\partial t_2}\right] = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right] \begin{bmatrix} \frac{\partial x}{\partial t_1} & \frac{\partial x}{\partial t_2} \\ \frac{\partial y}{\partial t_1} & \frac{\partial y}{\partial t_2} \\ \frac{\partial z}{\partial t_1} & \frac{\partial z}{\partial t_2} \end{bmatrix}$$

Therefore: $\dfrac{\partial f}{\partial t_1} = \dfrac{\partial f}{\partial x} \cdot \dfrac{\partial x}{\partial t_1} + \dfrac{\partial f}{\partial y} \dfrac{\partial y}{\partial t_1} + \dfrac{\partial f}{\partial z} \dfrac{\partial z}{\partial t_1}$.

$$\frac{\partial f}{\partial t_2} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t_2} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t_2} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t_2}.$$

(This is what is on usual multi-variable calculus book).

Module 3. Mean Value Theorem

Remark 7.3.1.
Recall for single-valued functions:
$f[a,b] \to \mathbb{R}$ is continuous. If f is differentiable on (a,b), then there exists $c \in (a,b)$ such that $f'(c) = \frac{f(b)-f(a)}{b-a}$.

Remark 7.3.2: Naive approach.
$U \in \mathbb{R}^n$ is open, $f: U \to \mathbb{R}^m$ is differentiable. If $a, b \in U$, then there exists $c \in L[a,b]$
$= \{(1-t)a + tb : t \in [0,1]\}$
such that

$$f(a) - f(b) = Df(c)(b-a).$$

Example 7.3.3: The above is false, unfortunately.
$f: \mathbb{R} \to \mathbb{R}^2$.
$f(x) = (\cos x, \sin x)$.
$f(0) = f(2\pi) = (1, 0)$.
But, $Df(x) = \begin{bmatrix} -\sin x \\ \cos x \end{bmatrix} \neq 0$.

So, $0 \neq 2\pi(-\sin x, \cos x)^T$
for all $x \in \mathbb{R}$.

Remark 7.3.4: Idea.
We will work "one direction at a time".

for all $x \in \mathbb{R}^m$, find $c \in L(a,b)$ such that

$$x \cdot (f(b) - f(a) - Df(c)(b-a)) = 0.$$

$$\Leftrightarrow \quad x \cdot (f(b) - f(a)) = x \cdot (Df(c)(b-a))$$

Lemma 7.3.5: $a, b \in \mathbb{R}^n$.

The function $\varphi : \mathbb{R} \to \mathbb{R}^n$

$$\varphi(t) = (1-t)a + tb$$

is differentiable with

$$D\varphi(t) = b - a.$$

Proof:

$$\lim_{h \to 0} \frac{\varphi(t+h) - \varphi(t) - (b-a)h}{\|h\|}$$

$$= \lim_{h \to 0} \frac{(1-t-h)a + (t+h)b - (1-t)a - tb - (b-a)h}{\|h\|}$$

$$= \lim_{h \to 0} \frac{-ha + hb - (b-a)h}{\|h\|}$$

$$= 0.$$

Theorem 7.3.6 [Mean Value Theorem].

Let $U \subseteq \mathbb{R}^n$ be open. If $f : U \to \mathbb{R}^m$ is differentiable and $a, b \in U$ such that $L(a,b) \subseteq U$, then for all $x \in \mathbb{R}^m$. there exists $c \in L(a,b)$ such that

$$x \cdot (f(b) - f(a)) = x \cdot (Df(c)(b-a)).$$

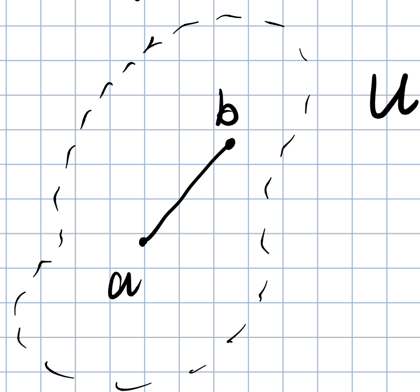
(c depends on the x chosen).

Proof: Let $L(a,b) \subseteq U$, $x \in \mathbb{R}^m$

① $\varphi(t) = (1-t)a + tb$.

$\varphi([0,1]) = L(a,b) \subseteq U$ (U is open).

$\Rightarrow \exists\, \delta > 0$ such that $\varphi(0-\delta, 1+\delta) \subseteq U$.



② For $t \in (-\delta, 1+\delta)$,

$$D(f\circ\varphi)(t) = Df(\varphi(t))(b-a)$$

③. $F : (-\delta, 1+\delta) \to \mathbb{R}$

$$F(t) = x \cdot (f\circ\varphi)(t).$$

By Dot product rule:

$$F'(t) = x \cdot Df(\varphi(t))(b-a).$$

④ By single value mean value theorem.

$\exists\, t_0 \in (0,1)$ such that

$F(1) - F(0) = F'(t_0)(1-0)$

$\Rightarrow x \cdot f(b) - x \cdot f(a) = x \cdot Df(\varphi(t_0))(b-a)$.

$\Rightarrow x \cdot (f(b) - f(a)) = x \cdot Df(\varphi(t_0))(b-a)$.

Choose $c = \varphi(t_0)$, we are done. $\square$

Module 4. Tangent Hyperplanes.

Remark 7.4.1: Goal

Understand the geometrical interpretation of the total derivative (gradient) of scalar functions:

$$f: \mathcal{U} \to \mathbb{R}.$$

where $\mathcal{U} \subseteq \mathbb{R}^n$ is open.

Remark 7.4.2: Motivation.

$n=1$, Let $\mathcal{U} \subseteq \mathbb{R}$ be open.

If $f: \mathcal{U} \to \mathbb{R}$ is differentiable at $a \in \mathcal{U}$, then $f'(a) = \nabla f(a)$ is the slope of the tangent line to the curve $y=f(x)$ at $x=a$.

When $n=2$, $\mathcal{U} \subseteq \mathbb{R}^2$ is open, we want:

If $f: \mathcal{U} \to \mathbb{R}$ is differentiable at $a \in \mathcal{U}$, then $Df(a) = \nabla f(a)$ tells us information about the tangent plane to the surface $z=f(x,y)$ at $(x,y)=a$.

Definition 7.4.3.

A hyperplane in $\mathbb{R}^n$ is a set of the form.

$$P = \{(x_1, x_2, \dots, x_n) \subseteq \mathbb{R}^n : a_1x_1 + a_2x_2 + \dots + a_nx_n = d\}$$

for some fixed $a_1, a_2, \dots, a_n \in \mathbb{R}$ (not all zero) and $d \in \mathbb{R}$.

Remark 7.4.4.

$n=2$, hyperplanes are lines.

$n=3$, hyperplanes are planes.

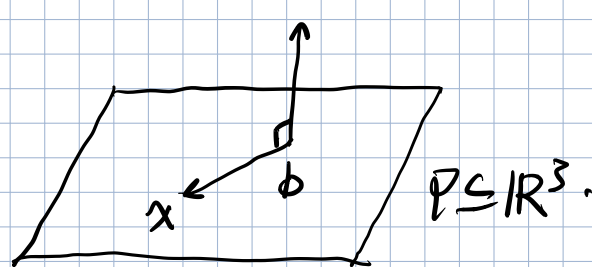
Example 7.4.5: $P=\{(x,y,z)\in\mathbb{R}^3: 2x+y-3z=1\}$.

Definition 7.4.6.
Let $P=\{(x_1,\dots,x_n): a_1x_1+\dots+a_nx_n=d\}$ be a hyperplane in $\mathbb{R}^n$. We call $n=(a_1,a_2,\dots,a_n)$ the normal (vector) of P.

Remark 7.4.7: Geometrically:
Let $b=(b_1,b_2,\dots,b_n)\in P$.
Then, $d=a_1b_1+a_2b_2+\dots+a_nb_n$
Therefore, $x=(x_1,\dots,x_n)\in P$

$$\begin{aligned}\Leftrightarrow\ & d=a_1x_1+a_2x_2+\dots+a_nx_n.\\ \Leftrightarrow\ & 0=d-d\\ &\ =a_1(x_1-b_1)+\dots+a_n(x_n-b_n)\\ &\ =n\cdot(x-b).\end{aligned}$$

So, $P=\{x\in\mathbb{R}^n: n\cdot(x-b)=0\}$.
ie. $x\in P$ iff n is orthogonal/prependicular to $x-b$.

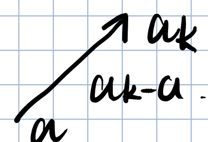


Definition 7.4.8: $A\subseteq\mathbb{R}^n$, $a\in A$.
A hyperplane $a\in P\subseteq\mathbb{R}^n$ with normal n is said to be tangent to A at a if

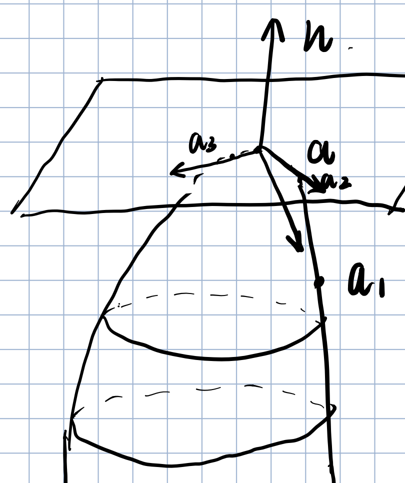
$$n \cdot \frac{a_k - a}{\|a_k - a\|} \to 0.$$

for all sequences $(a_k) \subseteq A \setminus \{a\}$ such that $a_k \to a$.

Remark 7.4.9: Why is this a good definition?
Recall that $a, b \in \mathbb{R}^n$ are orthogonal iff $a \cdot b = 0$.
Then, $n \cdot \frac{a_k - a}{\|a_k - a\|} \to 0$ says that unit (length 1) vectors in the direction of $a_k - a$:



become closer and closer to being orthogonal to n as $k \to \infty$.



Theorem 7.4.10:
$U \subseteq \mathbb{R}^n$ open, $a \in U$, $f: U \to \mathbb{R}$.
If f is differentiable at a, then the surface

$$S = \{(x, z) \in \mathbb{R}^{n+1} : z = f(x),\ x \in U\}.$$

has a tangent hyperplane at $(a, f(a))$ with the normal

$$n = (\nabla f(a), -1)$$

Proof: Appendix.

Example 7.4.11:

Find the tangent plane to the surface

$$z = 2x^2 + y^2.$$

at $(1, 1, 3)$.

Consider $f: \mathbb{R}^2 \to \mathbb{R}$

$$f(x,y) = 2x^2 + y^2.$$

Note f_x, f_y exist and are continuous on $\mathbb{R}^2$.

So, f is differentiable on $\mathbb{R}^2$.

$$\nabla f(x,y) = (4x, 2y).$$

$$\nabla f(1,1) = (4, 2).$$

The normal is:

$$n = (4, 2, -1).$$

The equation of the tangent plane is:

$$4x + 2y - z = d.$$

$$\Rightarrow d = 4\times 1 + 2\times 1 - 3 = 3.$$

$$\Rightarrow 4x + 2y - z = 3.$$

Week 7 Appendix

Theorem (Dot Product Rule)**.** Let $A \subseteq \mathbb{R}^n$ and let f and g be functions from A to $\mathbb{R}^m$. If f and g are differentiable at $a \in A$ then $f \cdot g$ is differentiable at a and

$$D(f \cdot g)(a) = g(a)Df(a) + f(a)Dg(a).$$

Proof. We must prove that

$$\lim_{h \to 0} \frac{(f \cdot g)(a+h) - (f \cdot g)(a) - Xh}{||h||} = 0,$$

where $X = g(a)Df(a) + f(a)Dg(a)$. To ease notations, let

$$\varepsilon(h) = f(a+h) - f(a) - Df(a)h$$

and

$$\delta(h) = g(a+h) - g(a) - Dg(a)h.$$

Since f and g are differentiable at a we have that

$$\lim_{h \to 0} \frac{\varepsilon(h)}{||h||} = \lim_{h \to 0} \frac{\delta(h)}{||h||} = 0.$$

Then,

$$\begin{aligned}
&\lim_{h \to 0} \frac{(f \cdot g)(a+h) - (f \cdot g)(a) - Xh}{||h||} \\
&= \lim_{h \to 0} \frac{(f \cdot g)(a+h) - (f \cdot g)(a) - g(a)Df(a)h - f(a)Dg(a)h}{||h||} \\
&= \lim_{h \to 0} \frac{g(a) \cdot \varepsilon(h) + f(a) \cdot \delta(h)}{||h||} \\
&+ \lim_{h \to 0} \frac{f(a) \cdot g(a) - g(a) \cdot f(a+h) - f(a) \cdot g(a+h) + f(a+h) \cdot g(a+h)}{||h||} \\
&= 0 + \lim_{h \to 0} \frac{f(a) \cdot g(a) - g(a) \cdot f(a+h) - f(a) \cdot g(a+h) + f(a+h) \cdot g(a+h)}{||h||} \\
&= \lim_{h \to 0} \frac{g(a) \cdot (f(a) - f(a+h)) - g(a+h) \cdot (f(a) - f(a+h))}{||h||} \\
&= \lim_{h \to 0} \frac{(g(a) - g(a+h)) \cdot (f(a) - f(a+h))}{||h||}.
\end{aligned}$$

However, by the Cauchy-Schwarz inequality,

$$\frac{|(g(a)-g(a+h))\cdot(f(a)-f(a+h))|}{||h||} \leq \frac{||g(a)-g(a+h)||\cdot||f(a)-f(a+h)||}{||h||}.$$

Therefore,

$$\begin{aligned}
0 &\leq \lim_{h\to 0}\frac{|(f\cdot g)(a+h)-(f\cdot g)(a)-Xh|}{||h||} \\
&\leq \lim_{h\to 0}\frac{||g(a)-g(a+h)||\cdot||f(a)-f(a+h)||}{||h||} \\
&= \lim_{h\to 0}\frac{||g(a)-g(a+h)||}{||h||}\cdot\frac{||f(a)-f(a+h)||}{||h||}||h|| \\
&= \lim_{h\to 0}\frac{||g(a)-g(a+h)-Dg(a)h+Dg(a)h||}{||h||}\cdot\frac{||f(a)-f(a+h)-Df(a)h+Df(a)h||}{||h||}||h|| \\
&\leq \lim_{h\to 0}\frac{||g(a)-g(a+h)-Dg(a)h||+||Dg(a)h||}{||h||}\cdot\frac{||f(a)-f(a+h)-Df(a)h||+||Df(a)h||}{||h||}||h|| \\
&\leq \lim_{h\to 0}\frac{||g(a)-g(a+h)-Dg(a)h||+||Dg(a)||_{op}||h||}{||h||}\cdot\frac{||f(a)-f(a+h)-Df(a)h||+||Df(a)||_{op}||h||}{||h||}||h|| \\
&= \lim_{h\to 0}(0+||Dg(a)||_{op})(0+||Df(a)||_{op})||h|| = 0
\end{aligned}$$

The result follows. $\square$

Theorem (Chain Rule)**.** Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ and consider two functions $f : A \to \mathbb{R}^m$ and $g : B \to \mathbb{R}^k$. If f is differentiable at $a \in A$ and g is differentiable at $f(a) \in B$ then $g \circ f$ is differentiable at a wth

$$D(g \circ f)(a) = Dg(f(a))Df(a).$$

Proof. We must prove that

$$\lim_{h \to 0} \frac{(g \circ f)(a+h) - (g \circ f)(a) - Xh}{||h||} = 0,$$

where $X = Dg(f(a))Df(a)$. To ease notation, let $b = f(a)$,

$$\varepsilon(h) = f(a+h) - f(a) - Df(a)h,$$

$$\delta(k) = g(b+k) - g(b) - Dg(b)k,$$

so that

$$\lim_{h \to 0} \frac{\varepsilon(h)}{||h||} = 0$$

and

$$\lim_{k \to 0} \frac{\delta(k)}{||k||} = 0.$$

Now, consider $k = f(a+h) - f(a)$. Note that $k \to 0$ as $h \to 0$, by continuity of f at a. Then,

$$\begin{aligned}
&\lim_{h \to 0} \frac{(g \circ f)(a+h) - (g \circ f)(a) - Xh}{||h||} \\
&= \lim_{h \to 0} \frac{(g \circ f)(a+h) - (g \circ f)(a) - Dg(f(a))Df(a)h}{||h||} \\
&= \lim_{h \to 0} \frac{g(k+b) - g(b) - Dg(b)Df(a)h}{||h||} \\
&= \lim_{h \to 0} \frac{Dg(b)\varepsilon(h) + \delta(k)}{||h||} \\
&= \lim_{h \to 0} Dg(b)\frac{\varepsilon(h)}{||h||} + \frac{\delta(k)}{||h||}.
\end{aligned}$$

Now, since

$$0 \leq \frac{||Dg(b)\varepsilon(h)||}{||h||} \leq ||Dg(b)||_{op} \frac{|||\varepsilon(h)||}{||h||} \to 0,$$

as $h \to 0$ we see that

$$\lim_{h \to 0} Dg(b)\frac{\varepsilon(h)}{||h||} = 0.$$

Next,
$$\lim_{h\to 0}\frac{\delta(k)}{||h||}=\lim_{h\to 0}\frac{\delta(k)}{||k||}\cdot\frac{||k||}{||h||}.$$
However,
$$||k||=||Df(a)h+\varepsilon(h)||\leq ||Df(a)||_{op}||h||+||\varepsilon(h)||,$$
from which it follows that
$$\frac{||k||}{||h||}$$
is bounded. By a squeeze theorem argument,
$$\lim_{h\to 0}\frac{\delta(k))}{||h||}=\lim_{h\to 0}\frac{\delta(k)}{||k||}\cdot\frac{||k||}{||h||}=0,$$
as required.

$\square$

Theorem. Let a be an element of an open set $U \subseteq \mathbb{R}^n$ and let $f : U \to \mathbb{R}$ be a function. If f is differentiable at a then the surface

$$S = \{(x, z) \in \mathbb{R}^{n+1} : z = f(x), x \in U\},$$

has a tangent hyperplane at $(a, f(a))$ with normal $n = (\nabla f(a), -1)$.

Proof. Let $(x_k, f(x_k)) \subseteq S \backslash \{(a, f(a))\}$ be a sequence such that $(x_k, f(x_k)) \to (a, f(a))$. By A1, $x_k \to a$. We must prove that

$$\lim_{k\to\infty} n \cdot \frac{(x_k, f(x_k)) - (a, f(a))}{||(x_k, f(x_k)) - (a, f(a))||} = 0.$$

Since f is differentiable at a we have that

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - \nabla f(a)h}{||h||} = 0.$$

Letting $\varepsilon(h) = f(a+h) - f(a) - \nabla f(a)h$,

$$\lim_{h\to 0} \frac{\varepsilon(h)}{||h||} = 0.$$

Moreover, we see that

$$||(x_k, f(x_k)) - (a, f(a))||^2 = ||(x_k - a, f(x_k) - f(a))||^2 \geq ||x_k - a||^2.$$

Then, since $x_k - a \to 0$,

$$\begin{aligned}
0 &\leq \left| \lim_{k\to\infty} n \cdot \frac{(x_k, f(x_k)) - (a, f(a))}{||(x_k, f(x_k)) - (a, f(a))||} \right| \\
&= \lim_{k\to\infty} \frac{|\nabla f(a)(x_k - a) - (f(x_k) - f(a))|}{||(x_k, f(x_k)) - (a, f(a))||} \\
&\leq \lim_{k\to\infty} \frac{|\nabla f(a)(x_k - a) - (f(x_k) - f(a))|}{||x_k - a||} \\
&= \lim_{k\to\infty} \frac{|\varepsilon(x_k - a)|}{||x_k - a||} \\
&= 0.
\end{aligned}$$

The result follows.

□

Week 8 Notes.

Module 1 Higher Order Total Derivatives.

Remark 8.1.1: Goal.

Come up with a sensible definition of an n^{th} order total derivative of a scalar function.

Remark 8.1.2: Idea.

Use n^{th} order partial derivative.

Remark 8.1.3: Notation.

$f: U \to \mathbb{R}$ is differentiable at $a \in U$. We may think of $\nabla f(a)$ as a function via matrix multiplication:

$$\nabla f(a): \mathbb{R}^n \to \mathbb{R}.$$

$$\nabla f(a)(h_1, h_2, \ldots, h_n) = \frac{\partial f}{\partial x_1}(a) h_1 + \cdots + \frac{\partial f}{\partial x_n}(a) h_n.$$

Definition 8.1.4:

Let $U \subseteq \mathbb{R}^n$ be open. $f: U \to \mathbb{R}$, $k \in \mathbb{N}$. Assume all partials of order less or equal to k exist at $a \subseteq U$. We define the k^{th} order total derivative of f at a by:

$$D^k f(a): \mathbb{R}^n \to \mathbb{R},$$

$$D^k f(a)(h_1, \ldots, h_n) = \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(a) h_{i_1} h_{i_2} \cdots h_{i_k}.$$

Example 8.1.5: $f: \mathbb{R}^2 \to \mathbb{R}$.

$$D^2 f(a)(h_1, h_2) = f_{xx}(a) h_1^2 + f_{xy}(a) h_1 h_2 + f_{yx}(a) h_2 h_1 + f_{yy}(a) h_2^2.$$

Module 8.2: Taylor's Theorem

The purpose of this short, written module is to prove that there is a multivariable analogue of Taylor's theorem. This result allows us to approximate a function, under certain circumstances, using higher order total derivatives and Taylor-like series.

Theorem. (Taylor's Theorem) Let $p \in \mathbb{N}$, $U \subseteq \mathbb{R}^n$ be open, and $f \in C^p(U)$. For all $x, a \in U$ with $L(x,a) \subseteq U$, there exists $c \in L(x,a)$ such that

$$f(x) = f(a) + \sum_{k=1}^{p-1} \frac{1}{k!} D^k f(a)(x-a) + \frac{1}{p!} D^p f(c)(x-a).$$

Proof. Let $x, a \in U$ and consider $h = x - a = (h_1, \ldots, h_n)$. Since $L(x,a) \subseteq U$ and U is open, there exists $\delta > 0$ such that $a + th \in U$ for all $t \in I := (-\delta, 1+\delta)$. Now, the function $g : I \to \mathbb{R}$ given by $g(t) = f(a+th)$ is differentiable by the chain rule and

$$g'(t) = Df(a+th)h = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a+th)h_i.$$

Moreover, it may be shown by induction that for $1 \leq j \leq p$,

$$g^{(j)}(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f}{\partial x_{i_1} x \partial x_{i_2} \cdots \partial x_{i_j}}(a+th)h_{i_1} \cdots h_{i_j}.$$

Note that this is the motivation for the definition for the higher-order total differivative! In particular, for $1 \leq j \leq p-1$ we have that

$$g^{(j)}(0) = D^j f(a)(h)$$

and

$$g^{(p)}(t) = D^p f(a+th)(h).$$

Therefore $g : I \to \mathbb{R}$ is p-times differentiable and so by the 1D version of Taylor's Theorem,

$$g(1) - g(0) = \sum_{k=1}^{p-1} \frac{1}{k!} g^{(k)}(0) + \frac{1}{p!} g^{(p)}(t),$$

for some $0 \leq t \leq 1$. Thus,

$$f(x) - f(a) = f(a+h) - f(a) = \sum_{k=1}^{p-1} \frac{1}{k!} D^k f(a)(h) + \frac{1}{p!} D^p f(a+th)(h),$$

and so we are done by taking $c = a + th$.

$\square$

Module 3. Optimization.

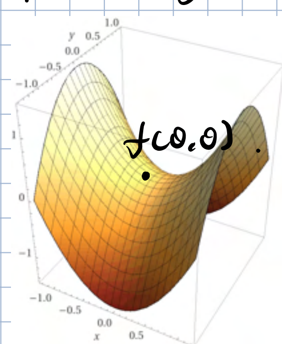
Definition 8.3.1.
Let $U \subseteq \mathbb{R}^n$ be open, $f: U \to \mathbb{R}$. Let $a \in U$.
① $f(a)$ is a local maxima (resp. local minima) of f if there exists $r > 0$ such that $f(x) \leq f(a)$ (resp. $f(x) \geq f(a)$) for all $x \in B_r(a)$.
② $f(a)$ is a local extrema if it is a local minima or maxima.

Remark 8.3.2.
Let U be open. $f: U \to \mathbb{R}$ is differentiable. $a \in U$.
Assume $f(a)$ is a local extrema of f, say $a = (a_1, \dots, a_n)$. Then

$$g_i(t) = f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n).$$

has a local extrema at $t = a_i$. Hence $g_i'(a_i) = 0$.
i.e. $\frac{\partial f}{\partial x_i}(a) = 0$. This implies $\nabla f(a) = 0$.

Example 8.3.3.
$f(x,y) = x^2 + y^2$. $\nabla f(x,y) = (2x, -2y)$.
$\Rightarrow \nabla f(0,0) = (0,0)$. But $f(0,0)$ is not an extrema.



This is something called a saddle point. A saddle point is a point where the gradient is zero but neither a maxima nor a minima.

Theorem 8.3.4 [2^{nd} Derivative Test].
Let $U \subseteq \mathbb{R}^n$ be open, $f \in C^2(U)$, $a \in U$. If $\nabla f(a) = 0$, then:

① $\forall h \neq 0$, $D^2 f(a)(h) > 0 \Rightarrow f(a)$ is a local minima.
② $\forall h \neq 0$, $D^2 f(a)(h) < 0 \Rightarrow f(a)$ is a local maxima.
③ $\exists h, k \in \mathbb{R}^n$ such that.

$$D^2 f(a)(h) > 0, \quad D^2 f(a)(k) < 0.$$

$\Rightarrow a$ is a saddle point.
Proof: Appendix

Lemma 8.3.5: $a, b, c \in \mathbb{R}$, $D := b^2 - ac$.
Consider $\varphi : \mathbb{R}^2 \to \mathbb{R}$.

$$\varphi(h,k) = ah^2 + 2bhk + ck^2.$$

① If $D < 0$, then a, $\varphi(h,k)$ have the same sign for all $(h,k) \neq 0$.
② If $D > 0$, then $\varphi(h,k)$ takes on both positive and negative values.
Proof: Omitted.

Theorem 8.3.6.
Let $U \subseteq \mathbb{R}^2$ be open, $f \in C^2(U)$, $\nabla f(a,b) = 0$.
$D := f_{xy}(a,b)^2 - f_{xx}(a,b) f_{yy}(a,b)$

(called the discriminant).

① If $D<0$ and $f_{xx}(a,b)>0$, then $f(a,b)$ is a local minima.

② If $D<0$ and $f_{xx}(a,b)<0$, then $f(a,b)$ is a local maxima.

③ If $D>0$, then (a,b) is a saddle point.

Proof Sketch.

Let $a=f_{xx}(a,b)$, $b=f_{xy}(a,b)$, $c=f_{yy}(a,b)$

$\varphi(h,k)=D^2f(a)(h,k)$ in the Lemma.

Module 4: Examples

Example 8.4.1.

Classify all local extrema and/or saddle points of $f(x,y) = x^4 + y^4 - 4xy + 2$.

Solution:

$Df(x,y) = (4x^3 - 4y,\ 4y^3 - 4x)$.

$Df(x,y) = 0 \Leftrightarrow \begin{cases} x^3 - y = 0 \\ y^3 - x = 0 \end{cases} \Leftrightarrow x = y^3,\ y = x^3$

$\Leftrightarrow x = x^9 \Leftrightarrow x = 0, y = 0$ or $x = 1, 1$. or $x = -1, y = -1$.

These are the critical points.

$f_{xx}(x,y) = 12x^2$, $f_{yy}(x,y) = 12y^2$, $f(x,y) = -4$.

① $(0,0)$.

$D = 16 - 0 > 0$. This is a saddle point.

② $(1,1)$.

$D = 16 - 12 \cdot 12 < 0$. $f_{xx}(1,1) = 12 > 0$.

$\Rightarrow f(1,1) = 0$ is a local minima.

③ $(-1,-1)$.

$D = 16 - 12 \cdot 12 < 0$. $f_{xx}(-1,-1) = 12 > 0$.

$\Rightarrow f(-1,-1) = 0$ is a local minima.

Example 8.4.2. $K = \overline{B_1(0,0)}$.

Find the absolute maxima and minima of $f: K \to \mathbb{R}$.

$$f(x,y) = 2x^3 + y^4.$$

ie. $\max f(K)$, $\min f(K)$.

① Critical points of $f: B_1(0,0) \to \mathbb{R}$.

$$\nabla f(x,y) = (6x^2, 4y^3) = (0,0)$$

$\Leftrightarrow (x,y) = (0,0)$. Note $f(0,0) = 0$.

② $\partial(K) = \{(x,y) : x^2+y^2=1\}$.

For $(x,y) \in \partial(K)$:

$$f(x,y) = 2x^3 + (1-x^2)^2 = \underbrace{x^4+2x^3-2x^2+1}_{g(x)}.$$

Consider $g(x)$ on $[-1,1]$.

$$g'(x) = 4x^3+6x^2-4x$$
$$= 2x(2x-1)(x+2) = 0$$

$\Leftrightarrow x = 0, \frac{1}{2}, -2$ (not in the domain).

Note $g(0)=1$, $g(\frac{1}{2}) = \frac{13}{16}$. $g(1)=2$, $g(-1)=-2$.

So, absolute maxima $f(1,0)=2$

absolute maxima $f(-1,0) = -2$.

Week 8 Appendix

Lemma. Let $U \subseteq \mathbb{R}^n$ be open and let $f \in C^2(U)$. If $a \in U$ such that $D^2f(a)(h) > 0$ for all $0 \neq h \in \mathbb{R}^n$ then there exists $m > 0$ such that

$$D^2f(a)(x) \geq m||x||^2,$$

for all $x \in \mathbb{R}^n$.

Proof. Consider the compact set $K = \{x \in \mathbb{R}^n : ||x|| = 1\}$. Since $f \in C^2(U)$ we have that $D^2f(a)$ is continuous and positive on K. By the EVT, there exists $m > 0$ such that $m = \min\{D^2f(a)(x) : x \in K\}$. For $0 \neq x \in \mathbb{R}^n$ we then see that $\frac{x}{||x||} \in K$ and so

$$D^2f(a)\left(\frac{x}{||x||}\right) = \frac{1}{||x||^2}D^2f(a)(x) \geq m.$$

□

Lemma. Let $U \subseteq \mathbb{R}^n$ be open and let $f \in C^2(U)$. Suppose $a \in U$ such that that $\nabla f(a) = 0$. Let $r > 0$ such that $B_r(a) \subseteq U$. There exists a function $\varepsilon : B_r(0) \to \mathbb{R}$ such that

$$\lim_{h \to 0} \varepsilon(h) = 0$$

and

$$f(a+h) - f(a) = \frac{1}{2}D^2f(a)(h) + ||h||^2\varepsilon(h)$$

for $||h||$ sufficiently small.

Proof. Consider

$$\varepsilon(h) := \frac{f(a+h) - f(a) - \frac{1}{2}D^2f(a)(h)}{||h||^2}$$

for $0 \neq h \in B_r(0)$ and define $\varepsilon(0) = 0$. We are left to prove that $\varepsilon(h) \to 0$ as $h \to 0$. Let $h \in B_r(0)$. Since $\nabla(f)(a) = 0$ we have by Taylor's Theorem that

$$f(a+h) - f(a) = \frac{1}{2}D^2f(c)(h)$$

for some $c \in L(a, a+h)$. Then,

$$\begin{aligned}
0 \leq |\varepsilon(h)|||h||^2 &= |\frac{1}{2}D^2f(c)(h) - \frac{1}{2}D^2f(a)(h)| \\
&\leq \frac{1}{2}\sum_i\sum_j \left|\frac{\partial^2 f}{\partial x_i \partial x_j}(c) - \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right| |h_i h_j| \\
&\leq \frac{1}{2}\sum_i\sum_j \left|\frac{\partial^2 f}{\partial x_i \partial x_j}(c) - \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right| ||h||^2
\end{aligned}$$

and

$$\frac{1}{2}\sum_i\sum_j\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(c) - \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right) \to 0$$

as $h \to 0$ because $c \to a$ as $h \to 0$ and $f \in C^2(U)$. □

Theorem. (Second Derivative Test) Let $U \subseteq \mathbb{R}^n$ be open and let $f \in C^2(U)$. Suppose $a \in U$ such that that $\nabla f(a) = 0$.

1. If $D^2 f(a)(h) > 0$ for all $0 \neq h \in \mathbb{R}^n$ then $f(a)$ is a local minimum of f.
2. If $D^2 f(a)(h) < 0$ for all $0 \neq h \in \mathbb{R}^n$ then $f(a)$ is a local maximum of f.
3. If there exist $h, k \in \mathbb{R}^n$ such that $D^2 f(a)(h) > 0$ and $D^2 f(a)(k) < 0$ then a is a saddle point of f.

Proof. Let $r > 0$ such that $B_r(a) \subseteq U$. There exists a function $\varepsilon : B_r(0) \to \mathbb{R}$ such that

$$\lim_{h \to 0} \varepsilon(h) = 0$$

and

$$f(a+h) - f(a) = \frac{1}{2} D^2 f(a)(h) + ||h||^2 \varepsilon(h)$$

for $||h||$ sufficiently small.

1. Suppose $D^2 f(a)(h) > 0$ for all $0 \neq h \in \mathbb{R}^n$. Let $m > 0$ such that
$$D^2 f(a)(x) \geq m||x||^2,$$
for all $x \in \mathbb{R}^n$. Then,
$$f(a+h) - f(a) = \frac{1}{2} D^2 f(a)(h) + ||h||^2 \varepsilon(h) \geq \left(\frac{m}{2} + \varepsilon(h)\right) ||h||^2 > 0$$
for all $||h||$ sufficiently small, since $m > 0$ and $\varepsilon(h) \to 0$ as $h \to 0$. Therefore $f(a+h) > f(a)$ for all $||h||$ sufficiently small, and so $f(a)$ is a local minimum of f.

2. Follows from (1) by replacing f with $-f$.

3. Let $h \in \mathbb{R}^n$. For small $t \in \mathbb{R}$,
$$\begin{aligned} f(a+th) - f(a) &= \frac{1}{2} D^2 f(a)(th) + ||th||^2 \varepsilon(th) \\ &= t^2 \left(\frac{1}{2} D^2 f(a)(h) + ||h||^2 \varepsilon(th)\right). \end{aligned}$$
Letting $t \to 0$, we have that $\varepsilon(th) \to 0$ and so $f(a+th) - f(a)$ takes on the same sign as $D^2 f(a)(h)$, which can be both positive and negative. Therefore a is a saddle point.

□

Week 9 Notes.

Module 1: Inverse Function Theorem.

Remark 9.1.1: Recall

$I=(a,b)$. If $f:I\to\mathbb{R}$ is continuous and injective, and $y\in f(I)$ is a point such that

1) f is differentiable at $x=f^{-1}(y)\in I$.
2) $f'(x)\neq 0$.

Then f^{-1} is differentiable at y and $(f^{-1})'(y)=\frac{1}{f'(x)}$.

Remark 9.1.2: Goal.

Develop a multivariable version of this theorem. We need to generalize the idea of $\frac{1}{f'(x)}=(f'(x))^{-1}$ to something like $Df(x)^{-1}$ (matrix inverse).

Definition 9.1.3:

Let $U\subseteq\mathbb{R}^n$, $f:U\to\mathbb{R}^n$. We define the Jacobian of f at a by

$$Jf(a):=\det(Df(a)).$$

Theorem 9.1.4 [Inverse Function Theorem].

Let $U\subseteq\mathbb{R}^n$ be open, $f\in C^1(U,\mathbb{R}^n)$. If $a\in U$ such that $Jf(a)\neq 0$, then there exists an open $a\in W\subseteq U$ such that.

1) f is injective on W.
2) $f^{-1}\in C^1(f(W),\mathbb{R}^n)$.
3) For all $y\in f(W)$

$$D(f^{-1})(y) = [Df(x)]^{-1}$$

where $x = f^{-1}(y)$.

Proof: Appendix. This is quite technical.

Example 9.1.5. $f: \mathbb{R}^2 \to \mathbb{R}^2$.

$$f(x,y) = (x+y, \sin x + \cos y).$$

Note that

$$f_x(x,y) = (1, \cos x).$$

$$f_y(x,y) = (1, -\sin y).$$

So that $f \in C^1(\mathbb{R}^2, \mathbb{R}^2)$.

Question: Prove that f^{-1} exists and is differentiable on some open set containing $(0,1)$ and compute $D(f^{-1})(0,1)$.

Note: $f(x,y) = (0,1)$ (Preimage).

$\Leftrightarrow (x+y, \sin x + \cos y) = (0,1)$

$\Leftrightarrow y = -x, \ \sin x + \cos(-x) = 1$.

$\Leftrightarrow y = -x, \ \sin x + \cos x = 1$.

$\Leftrightarrow (x,y) = (2k\pi, -2k\pi), \ k \in \mathbb{Z}$.

or $(x,y) = (\frac{\pi}{2} + 2k\pi, -\frac{\pi}{2} - 2k\pi), \ k \in \mathbb{Z}$

Case 1: $a = (2k\pi, -2k\pi), \ k \in \mathbb{Z}$.

$$Jf(a) = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1 \neq 0.$$

Then by the inverse function theorem, we have the results we need to prove.

Moreover,

$$D(f^{-1})(0,1) = [Df(a)]^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}.$$

Case 2: $a = (\frac{\pi}{2} + 2k\pi, -\frac{\pi}{2} - 2k\pi)$.

$$Jf(a) = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1 \neq 0.$$

Again, by the inverse function theorem, we have the results we need to prove. Moreover:

$$D(f^{-1})(0,1) = Df(a)^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

Remark 9.1.6:

The way we restrict f to make it injective depends on our choice of $f^{-1}(y)$.

Module 2: Implicit Function Theorem.

Remark 9.2.1: Idea.

When/where can $f(x,y,z)=0$ be solved to express z as a function of x,y?

That is when/where

$$\{(x,y,z)\in\mathbb{R}^3 : f(x,y,z)=0\}=\{(x,y,g(x,y)) : f(x,y,g(x,y))=0\}$$

Example 9.2.2.

$$f(x,y,z)=x^2+y^2+z^2-1=0.$$

$U=\{(x,y,z)\in\mathbb{R}^3 : z>0.\}$ is open.

On U, we have

$$z=\sqrt{1-x^2-y^2}:=g(x,y)$$

and $f(x,y,g(x,y))=0$.

Theorem 9.2.3 [Implicit Function Theorem].

Let $U\subseteq\mathbb{R}^{n+p}$ be open. $f=(f_1,\dots,f_n)\in C^1(U,\mathbb{R}^n)$

Let $x_0\in\mathbb{R}^n$, $t_0\in\mathbb{R}^p$ such that $f(x_0,t_0)=0$.

If

$$\det\left[\frac{\partial f_i}{\partial x_j}(x_0,t_0)\right]_{n\times n}\neq 0 \rightarrow \text{ignore other variables.}$$

then there exists an open $t_0\in V\subseteq\mathbb{R}^p$ and a unique $g\in C^1(V,\mathbb{R}^n)$ such that.

① $g(t_0)=x_0$ and

② $f(g(t),t)=0$, $\forall t\in V$.

Proof: Appendix.

Remark 9.2.4: Summary:

$t \in V \subseteq \mathbb{R}^p \to$ variables to keep.

$g(t) \in \mathbb{R}^n \to$ variables replaced by an implicit function of t.

Example 9.2.5: $xyz + \sin(x+y+z) = 0$.

Consider $f(x,y,z) = xyz + \sin(x+y+z)$

so that $f \in C^1(\mathbb{R}^3)$.

Note: $f(0,0,0) = 0$.

Now,

$$f_z(x,y,z) = xy + \cos(x+y+z)$$

$$\Rightarrow f_z(0,0,0) = 1 \neq 0.$$

$$\Rightarrow \det[1] = 1 \neq 0.$$

By the implicit function theorem, there exists open $V \subseteq \mathbb{R}^2$ with $(0,0) \in V$ and $g(x,y)$ in $C^1(V)$ such that $g(0,0) = 0$ and

$$f(x, y, g(x,y)) = 0.$$

for all $(x,y) \in V$. i.e. $z = g(x,y)$ on V.

Example 9.2.6.

Prove there exists $u, v : \mathbb{R}^4 \to \mathbb{R}$

and $(2,1,-1,-2) \in \mathcal{U} \subseteq \mathbb{R}^4$ open such that.

① $u, v \in C^1(\mathcal{U})$.

② $u(2,1,-1,-2) = 4$, $v(2,1,-1,-2) = 3$.

③ For all $(x,y,z,w) \in \mathcal{U}$

$$u^2 + v^2 + w^2 = 29.$$

$$\frac{u^2}{x^2}+\frac{v^2}{y^2}+\frac{w^2}{z^2}=17.$$

Solution:

$f: \mathbb{R}^6 \to \mathbb{R}^2$.

$f(u, v, x, y, z, w)$

$$= \left(u^2+v^2+w^2-29,\ \frac{u^2}{x^2}+\frac{v^2}{y^2}+\frac{w^2}{z^2}-17\right).$$

$$\Rightarrow f(4,3,2,1,-1,-2)=0.$$

and

$$\det\begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} \\ \frac{\partial f_3}{\partial u} & \frac{\partial f_2}{\partial v}\end{bmatrix} = \begin{vmatrix} 2u & 2v \\ \frac{2u}{x^2} & \frac{2v}{y^2}\end{vmatrix}.$$

$$=4uv\left(\frac{1}{y^2}-\frac{1}{x^2}\right).$$

This is non-zero at $(4,3,2,1,-1,-2)$

By the Implicit function theorem, there exists open $(2,1,-1,-2) \in U$ and $g \in C^1(U, \mathbb{R}^2)$ such that $g(2,1,-1,-2)=(4,3)$ and $\forall (x,y,z,w) \in U$,

$$f(g(x,y,z,w), x, y, z, w)=0.$$

Let $g=(g_1, g_2)$.

Take $u(x,y,z,w)=g_1(x,y,z,w)$

$v(x,y,z,w)=g_2(x,y,z,w)$.

$u, v \in C^1(U)$.

$$u(2,1,-1,-2)=4.$$

$$v(2,1,-1,-2)=3.$$

$$f(g(x,y,z,w), x, y, z, w)=0.$$

$$\Rightarrow f(u(x,y,z,w), v(x,y,z,w), x, y, z, w)=0.$$

$$\Rightarrow \quad u^2 + v^2 + w^2 = 29.$$

$$\frac{u^2}{x^2} + \frac{v^2}{y^2} + \frac{w^2}{z^2} = 17.$$

Module 3: Lagrange Multipliers.

Remark 9.3.1: Goal.

Solve optimization problems with constraints.

Definition 9.3.2:

Let $U \subseteq \mathbb{R}^n$ be open, $f: U \to \mathbb{R}$.

① Let $a \in U$. We call $f(a)$ a local max of f. Subject to the constraints

$$g_i: U \to \mathbb{R}.$$
$$1 \le i \le m.$$

If $g_i(a) = 0$ for all i and there exists $r > 0$ such that whenever $x \in B_r(a)$ and $g_i(x) = 0$ for all i, then $f(x) \le f(a)$.

②. Similar definition for local min.

Theorem 9.3.3.

Let $U \subseteq \mathbb{R}^n$ be open, $m < n$, $f, g_1, g_2, \dots, g_m \in C^1(U)$

Suppose $a \in U$ such that.

$$\det\left[\frac{\partial g_i}{\partial x_j}(a)\right]_{m \times m} \neq 0$$

If $f(a)$ is a local extremum of f subject to the constraints g_i, then $\exists\, \lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{R}$ such that

$$\nabla f(a) + \sum_{i=1}^{m} \lambda_i \nabla g_i(a) = 0.$$

Remark 9.3.4.

The proof for the above theorem is quite technical. We only consider the case $m=2, n=3$ here.

We want to show that there exists $\lambda, \mu \in \mathbb{R}$ such that

$$\lambda \frac{\partial g_1}{\partial x_j}(a) + \mu \frac{\partial g_2}{\partial x_j}(a) = -\frac{\partial f}{\partial x_j}(a).$$

for $j = 1, 2, 3$.

Let $A = \begin{bmatrix} \frac{\partial g_1}{\partial x_1}(a) & \frac{\partial g_1}{\partial x_2}(a) \\ \frac{\partial g_2}{\partial x_1}(a) & \frac{\partial g_2}{\partial x_2}(a) \end{bmatrix}$.

so that $\det A \neq 0$.

In particular

$$[\lambda \ \mu] A = \left[-\frac{\partial f}{\partial x_1}(a), -\frac{\partial f}{\partial x_2}(a)\right]$$

has a unique solution $[\lambda, \mu]$. Expand this we get the case for x_1, x_2. We are left to show that

$$\lambda \frac{\partial g_1}{\partial x_3}(a) + \mu \frac{\partial g_2}{\partial x_3}(a) = -\frac{\partial f}{\partial x_3}(a).$$

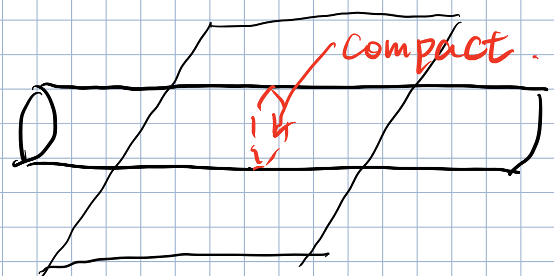
The idea is then to use the implicit function theorem to replace x_3 with $h(x_1, x_2)$ and prove the above result with what we have with x_1, x_2.

Example 9.3.5.

Maximize and minimize $f(x,y,z) = x+2y$, subject to the constraints

① $x+y+z=1$

② $y^2+z^2=4$.



Geometrically, such a max/min must exist!

Let $f(x,y,z) = x+2y$, $g(x,y,z) = x+y+z-1$

$h(x,y,z) = y^2+z^2-4$.

Note:

$$\begin{vmatrix} g_x & g_y \\ h_x & h_y \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 0 & 2y \end{vmatrix} = 2y \neq 0 \text{ for } y \neq 0.$$

Moreover, if $g(x,0,z) = h(x,0,z) = 0$

then $z = \pm\sqrt{2}$, $x = 1 \mp \sqrt{2}$. In which case.

$$f(1\mp\sqrt{2}, 0, \pm\sqrt{2}) = 1 \mp \sqrt{2}.$$

Otherwise, such a max/min is of the form

$$\nabla f = \lambda \nabla g + \mu \nabla h.$$

$\Rightarrow [1,2,0] = \lambda[1,1,1] + \mu[0,2y,2z]$.

$\Rightarrow \lambda = 1$, $\lambda + 2\mu y = 2$, $\lambda + 2\mu z = 0$.

$\Rightarrow y = \frac{1}{2\mu}$, $z = \frac{-1}{2\mu}$.

But $y^2+z^2=4 \Rightarrow \frac{2}{4\mu^2}=4$.

$\Rightarrow \mu=\pm\frac{1}{2\sqrt{2}}$.

$\Rightarrow y=\pm\sqrt{2},\ z=\mp\sqrt{2}$.

$\Rightarrow x=1$.

Checking:

$f(1,\sqrt{2},-\sqrt{2})=1+2\sqrt{2}$: max

$f(1,-\sqrt{2},\sqrt{2})=1-2\sqrt{2}$: min.

Week 9 Appendix

Lemma 1. Let $U \subseteq \mathbb{R}^n$ be open. Suppose $a \in U$ so that we may find $r > 0$ such that $\overline{B_r(a)} \subseteq U$. Let $f : U \to \mathbb{R}^n$ be continuous and injective when restricted to $\overline{B_r(a)}$ and assume its first order partials exist on $B_r(a)$. If $Jf \neq 0$ on $B_r(a)$ then there exists $\varepsilon > 0$ such that $B_\varepsilon(f(a)) \subseteq f(B_r(a))$.

Proof. Consider $g : \overline{B_r(a)} \to \mathbb{R}$ given by $g(x) = ||f(x) - f(a)||$. Since f is continuous and injective on $\overline{B_r(a)}$ we have that g is continuous and $g(x) > 0$ for all $x \neq a$. By the EVT,

$$m = \inf\{g(x) : ||x - a|| = r\} > 0.$$

Take $\varepsilon = m/2$. We claim that $B_\varepsilon(f(a)) \subseteq f(B_r(a))$.

Let $y \in B_\varepsilon(f(a))$. Again by the EVT, there exists $b \in \overline{B_r(a)}$ such that

$$||f(b) - y|| = \inf\{||f(x) - y|| : x \in \overline{B_r(a)}\}.$$

For the sake of contradiction, suppose that $||b - a|| = r$. Then,

$$\varepsilon > ||f(a)-y|| \geq ||f(b)-y|| \geq ||f(b)-f(a)||-||f(a)-y|| = g(b)-||f(a)-y|| \geq m-\varepsilon = 2\varepsilon-\varepsilon = \varepsilon,$$

which is a contradiction. Therefore we have that $b \in B_r(a)$.

If we can show that $y = f(b)$ we are done. This is where the information about the partial derivatives and the Jacobian come into play. Consider the continuous function $h : \overline{B_r(a)} \to \mathbb{R}$ given by $h(x) = ||f(x) - y||$. By construction, $h(b)$ is the minimum value of h. Moreover, $h^2(b)$ is also the minimum value of of h^2. Since $b \in B_r(a)$, which is open, we have that $\nabla h^2(b) = 0$ (note that in last week's proof we really just needed the first order partials to exist at a, not necessarily differentiability at a). However,

$$h^2(x) = \sum_{i=1}^{n} (f_i(x) - y_i)^2,$$

and so for every $1 \leq j \leq n$,

$$0 = \frac{\partial h^2}{\partial x_j}(b) = \sum_{i=1}^{n} 2(f_i(b) - y_i)\frac{\partial f_i}{\partial x_j}(b).$$

Thus, $Df(b)x = 0$, where $x = (2(f_1(b)-y_1), 2(f_2(b)-y_2), \ldots, 2(f_n(b)-y_n))^T$. Since $Df(b)$ is invertible ($Jf(b) \neq 0$) we have that $x = 0$. Hence $f(b) = y$, as required. $\square$

Lemma 2. Let $U \subseteq \mathbb{R}^n$ be open and nonempty. If $f : U \to \mathbb{R}^n$ is continuous, injective, has all first-order partials existing on U, AND is such that $Jf \neq 0$ on U, then f^{-1} is continuous on $f(U)$.

Proof. To prove that $f^{-1} : f(U) \to \mathbb{R}^n$ is continuous it suffices to prove that $f(W)$ is open whenever W is open in $\mathbb{R}^n$ and $W \subseteq U$ (Why? Piazza!). Well, let W be such a set and take $b \in f(W)$ so that $b = f(a)$ for some $a \in W$. Since W is open there exists $r > 0$ such that $\overline{B_r(a)} \subseteq W$. By the previous lemma, there then exists $\varepsilon > 0$ such that

$$B_\varepsilon(b) \subseteq f(B_r(a)).$$

Thus, $B_\varepsilon(b) \subseteq f(W)$, and so $f(W)$ is open. $\square$

Lemma 3. Let $U \subseteq \mathbb{R}^n$ be open and let $f \in C^1(U, \mathbb{R}^n)$. If $a \in U$ such that $Jf(a) \neq 0$ then there exists $r > 0$ such that $B_r(a) \subseteq U$, f is injective on $B_r(a)$, $Jf \neq 0$ on $B_r(a)$, and

$$\det\left(\frac{\partial f_i}{\partial x_j}(c_i)\right) \neq 0$$

for all $c_1, c_2, \ldots, c_n \in B_r(a)$.

Proof. Let $W = U \times U \times \cdots \times U$ (n-times). Consider $h : W \to \mathbb{R}$ defined by

$$h(x_1, x_2, \ldots, x_n) = \det\left(\frac{\partial f_i}{\partial x_j}(x_i)\right)$$

Since $f \in C^1(U, \mathbb{R}^n)$ and a determinant is a polynomial in its entries, we have that h is continuous. Note that $h(a, a, \ldots, a) = Jf(a) \neq 0$. Thus we may find an open interval $h(a, a, \ldots, a) \in I \subseteq \mathbb{R}$ such that $0 \notin I$. Then, $h^{-1}(I)$ is open (note that W is open) and so there exists $R > 0$ such that $B_R(a, a, \ldots, a) \subseteq h^{-1}(I)$. But then we may find $r > 0$ such that

$$B_r(a) \times \cdots \times B_r(a) \subseteq B_R(a, a, \ldots, a) \subseteq h^{-1}(I).$$

We then see that $Jf \neq 0$ on $B_r(a)$, and

$$\det\left(\frac{\partial f_i}{\partial x_j}(c_i)\right) \neq 0$$

for all $c_1, c_2, \ldots, c_n \in B_r(a)$.

We are left to show that f injective on $B_r(a)$. For the sake of contradiction suppose there exists $x \neq y$ in $B_r(a)$ such that $f(x) = f(y)$. Since f is differentiable on $B_r(a)$, every f_i is differentiable on $B_r(a)$. Fix $1 \leq i \leq n$. By the MVT there exists $c_i \in L(x, y)$ such that $0 = f_i(x) - f_i(y) = Df_i(c_i)(x-y)$. Letting $A = \left[\frac{\partial f_i}{\partial x_j}(c_i)\right]$ we see that $A(x-y) = 0$. Since $x - y \neq 0$, A is not invertible and so

$$\det\left(\frac{\partial f_i}{\partial x_j}(c_i)\right) = 0,$$

a contradiction.

□

Recall. (Cramer's Rule) Let A be a $n \times n$ invertible matrix and consider a system of equations $Ax = b$. This system has a unique solution $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ given by

$$x_i = \frac{\det(A(i))}{\det A},$$

where $A(i)$ is the matrix obtained from A by replacing its i^{th} column by b.

Theorem. (Inverse Function Theorem) Let $U \subseteq \mathbb{R}^n$ be open and let $f \in C^1(U, \mathbb{R}^n)$. If $a \in U$ such that $Jf(a) \neq 0$ then there exists an open set $a \in W \subseteq U$ such that

1. f is injective on W
2. $f^{-1} \in C^1(f(W), \mathbb{R}^n)$
3. For every $y \in f(W)$, if $x = f^{-1}(y)$ then
$$Df^{-1}(y) = [Df(x)]^{-1}.$$

Proof. Since this is a rather long and technical proof, we break it into digestible, enumerated pieces.

1. By Lemma 3 there exists $r > 0$ with $W := B_r(a) \subseteq U$ such that f is injective on W, $Jf \neq 0$ on W, and
$$\det\left(\frac{\partial f_i}{\partial x_j}(c_i)\right) \neq 0$$
for all $c_1, c_2, \dots, c_n \in W$. Moreover, by Lemma 2, f^{-1} is continuous on $f(W)$.

2. We claim that $f^{-1} \in C^1(f(W), \mathbb{R}^n)$. Fix $y_0 \in f(W)$ and $1 \leq i, j \leq n$. Choose $0 \neq t \in \mathbb{R}$ sufficiently small so that $y_0 + te_j \in f(W)$. We may then find $x_0, x_1 = x_1(t) \in W$ such that $f(x_0) = y_0$ and $f(x_1) = y_0 + te_j$. By the MVT, for every $1 \leq i \leq n$ there exists $c_i = c_i(t) \in L(x_0, x_1)$ such that
$$\nabla f_i(c_i)(x_1 - x_0) = f_i(x_1) - f_i(x_0) = \begin{cases} t & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

Therefore,

$$\nabla f_i(c_i)(\frac{x_1 - x_0}{t}) = \frac{1}{t}(f_i(x_1) - f_i(x_0)) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

Now let A_j be the $n \times n$ matrix whose i^{th} row is $\nabla f_i(c_i)$. By assumption, $\det(A_j) \neq 0$. Moreover, $A_j(\frac{x_1 - x_0}{t}) = e_j$. For $1 \leq k \leq n$, we then see that

$$\frac{(f^{-1})_k(y_0 + te_j) - (f^{-1})_k(y_0)}{t} = \frac{x_{1,k} - x_{0,k}}{t},$$

where by Cramer's Rule, $Q_k(t) := \frac{x_{1,k} - x_{0,k}}{t}$ is a quotient of determinants of matrices whose entries are either 0, 1, or a first-order partial of f evaluated at a c_ℓ. As $t \to 0$ we clearly have that $y_0 + te_j \to y_0$. But then, by the continuity of f^{-1}, we have that $x_1 \to x_0$ and so $c_i \to x_0$. Since f is C^1, we therefore that that $Q_k(t) \to Q_k$, where Q_k is a quotient of determinants of matrices whose entries are either 0, 1, or a first-order partial of f evaluated at a $x_0 = f^{-1}(y_0)$. Since $f \in C^1$ and f^{-1} is continuous at y_0, it follows that Q_k is continuous at each $y_0 \in f(W)$. Moreover,

$$\lim_{t \to 0} \frac{(f^{-1})_k(y_0 + te_j) - (f^{-1})_k(y_0)}{t} = \lim_{t \to 0} \frac{x_{1,k} - x_{0,k}}{t} = Q_k.$$

Hence all of the partial derivatives of f^{-1} exist and are continuous at y_0 (ie. $f^{-1} \in C^1(f(W), \mathbb{R}^n)$).

3. Finally, we quickly run the chain rule and note that for $y \in f(W)$,

$$I = DI(y) = D(f \circ f^{-1})(y) = Df(f^{-1}(y))D(f^{-1})(y).$$

The result follows.

□

Theorem. (Implicit Function Theorem) Suppose $U \subseteq \mathbb{R}^{n+p}$ is open and $f = (f_1, f_2, \dots, f_n) \in C^1(U, \mathbb{R}^n)$. Suppose $x_0 \in \mathbb{R}^n$ and $t_0 \in \mathbb{R}^p$ such that $f(x_0, t_0) = 0$. If

$$\det\left(\frac{\partial f_i}{\partial x_j}(x_0, t_0)\right)_{n \times n} \neq 0,$$

then there is an open set $t_0 \in V \subseteq \mathbb{R}^p$ and a unique function $g \in C^1(V, \mathbb{R}^n)$ such that $g(t_0) = x_0$ and $f(g(t), t) = 0$ for all $t \in V$.

Proof. For every $(x, t) \in U$ let

$$F(x, t) := (f(x, t), t) = (f_1(x, t), \dots, f_n(x, t), t_1, t_2, \dots, t_p).$$

Notice that $F(x_0, t_0) = (0, t_0)$. Then, F is a function from U to $\mathbb{R}^{n+p}$ with

$$DF = \begin{bmatrix} \left(\frac{\partial f_i}{\partial x_j}\right)_{n\times n} & B \\ 0_{p\times n} & I_{p\times p} \end{bmatrix},$$

where $0_{p\times n}$ is the $p\times n$ zero matrix, $I_{p\times p}$ is the $p\times p$ identity matrix, and B is a matrix whose entries are first-order partials of the $f_i's$ with respect to the $t_j's$. Taking the determinant of this crazy matrix evaluated at (x_0, t_0), we have that

$$JF(x_0, t_0) = \det\left(\frac{\partial f_i}{\partial x_j}(x_0, t_0)\right)_{n\times n} \cdot \det I_{n\times p} \neq 0.$$

Therefore, by the Inverse Function Theorem there exists an open set $(x_0, t_0) \in W \subseteq U$ such that F is injective on W and $F^{-1} \in C^1(F(W), \mathbb{R}^{n+p})$.

To ease notation, let $G = F^{-1} = (G_1, G_2, \ldots, G_n, G_{n+1}, \ldots, G_{n+p})$. Consider $\varphi : F(W) \to \mathbb{R}^n$ given by

$$\varphi = (G_1, G_2, \ldots, G_n).$$

By construction we have that

$$\varphi(F(x, t)) = x$$

for all $(x, t) \in W$ and

$$F(\varphi(x, t), t) = (x, t),$$

for al $(x, t) \in F(W)$.

Consider $V = \{t \in \mathbb{R}^p : (0, t) \in F(W)\}$ and the function $g : V \to \mathbb{R}^n$ given by $g(t) = \varphi(0, t)$. Since G is C^1, it follows that φ is also C^1. Hence, $g \in C^1(V, \mathbb{R}^n)$. Also note that V is open since $F(W)$ is open. Finally, we compute that

$$g(t_0) = \varphi(0, t_0) = \varphi(F(x_0, t_0)) = x_0,$$

and note that for all $(x, t) \in F(W)$,

$$f(\varphi(x, t), t) = x.$$

In particular,

$$0 = f(\varphi(0, t), t) = f(g(t), t) = 0$$

for all $t \in V$.

Uniqueness follows from the injectivity of F. (Please share the details on Piazza!)

□

Week 10 Notes.

Module 1: Riemann Integration

Remark 10.1.1: Recall $f: [a,b] \to \mathbb{R}$ bounded.

① A partition of $[a,b]$ is a set $P = \{x_0, x_1, \dots, x_n\}$ where $a = x_0 < x_1 < \cdots < x_n = b$.

② For a partion $P = \{x_0, \dots, x_n\}$:

$$U(f,P) = \sum_{i=1}^{n} M_i (x_i - x_{i-1}).$$

$$L(f,P) = \sum_{i=1}^{n} m_i (x_i - x_{i-1}).$$

where

$$m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}.$$

$$M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\}.$$

③ We say f is integrable iff

$$\inf_P \{U(f,P)\} = \sup_P \{L(f,P)\}.$$

$:= \overline{\int_a^b} f(x)\,dx$ Upper Riemann Integral

$:= \underline{\int_a^b} f(x)\,dx$ Lower Riemann Integral.

$\int_a^b f(x)\,dx$.

Definition 10.1.2:

A rectangle in $\mathbb{R}^n$ is a set of the form

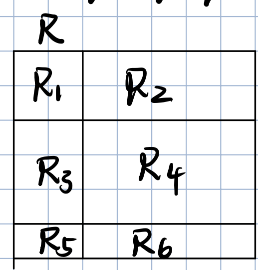
$$R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n].$$

where $a_i \le b_i$.

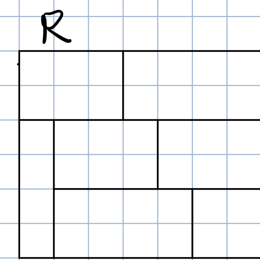
Definition 10.1.3:

A partition of $R=[a_1,b_1]\times\cdots\times[a_n b_n]$ is a grid of rectangles on R obtained by partitioning each $[a_i, b_i]$.

Example 10.1.4: In $\mathbb{R}^2$:



This is a partition



This is not a partition.

Definition 10.1.5

Let $R=[a_1,b_1]\times\cdots\times[a_n,b_n]\subseteq\mathbb{R}^n$ be a rectangle. The volume of R is:

$$V(R)=(b_1-a_1)(b_2-a_2)\cdots(b_n-a_n).$$

Definition 10.1.6:

Let $R\subseteq\mathbb{R}^n$ be a rectangle and let $f:R\to\mathbb{R}$ be bounded. Let $P=\{R_1,\ldots,R_n\}$ be a partition of R.

① Upper sum of f relative to P is:

$$U(f,P)=\sum_{i=1}^{n} M_i V(R_i).$$

where

$$M_i=\sup\{f(x): x\in R_i\}.$$

② Lower sum of f relative to P is:

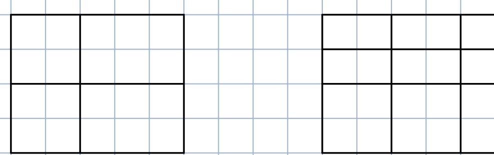
$$L(f,P)=\sum_{i=1}^{n} m_i V(R_i).$$

where

$$m_i=\inf\{f(x): x\in R_i\}.$$

Definition 10.1.7:
Let P, Q be partitions of $R \subseteq \mathbb{R}^n$. We say P is a refinement of Q, written $P \leq Q$, if P is obtained from Q by partitioning the sides of R even further.

Example 10.1.8: $P \leq Q$:



Q. $\qquad$ P

Proposition 10.1.9:
If $P \leq Q$ on R and $f: R \to \mathbb{R}$ is bounded, then

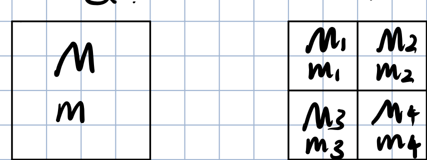
$$U(f,P) \leq U(f,Q)$$

and

$$L(f,P) \geq L(f,Q).$$

Proof sketch:

Q. $\qquad$ P



$$MV(R) = MV(R_1) + \cdots + MV(R_4)$$
$$\geq M_1 V(R_1) + \cdots + M_4 V(R_4)$$
$$m V(R) \leq m_1 V(R_1) + \cdots + m_4 V(R_4).$$

Definition 10.1.10
Let $f: R \to \mathbb{R}$ be bounded, where $R \subseteq \mathbb{R}^n$ be a rectangle.

① The lower integral of f is:

$$\underline{\int_R} f = \sup\{L(f,P) : P\}.$$

② The Upper integral of f is:

$$\overline{\int_R} f = \inf\{U(f,P) : P\}.$$

③ We say f is (Riemann) integrable over R if.

$$\int_R f := \underline{\int_R} f = \overline{\int_R} f.$$

Example 10.1.11: $f : [0,1] \times [0,1] \to \mathbb{R}$.

$$f(x,y) = \begin{cases} 1 & x \in \mathbb{Q}. \\ 0 & x \notin \mathbb{Q}. \end{cases}$$

for all P, we have $U(f,P) = 1$, $L(f,P) = 0$

So, $\underline{\int_R} f = 0 \neq 1 = \overline{\int_R} f$, f is not integrable.

Module 2: Characterization Theorem.

Remark 10.2.1: Goal

Give an alternative definition of $f: R \to \mathbb{R}$ being integrable.

Lemma 10.2.2:

Let $f: R \to \mathbb{R}$ be bounded. If P, Q are partitions of R, then $L(f,P) \leq U(f,Q)$.

Proof sketch:

Find a common refinement S (Eg overlap P and Q).

Then,
$$L(f,P) \leq L(f,S) \leq U(f,S) \leq U(f,Q).$$

Remark 10.2.3:

Let $f: R \to \mathbb{R}$ be bounded. For all partitions P, Q
$$L(f,P) \leq U(f,Q).$$
$$\Rightarrow \quad L(f,P) \leq \overline{\int_R} f$$
$$\Rightarrow \quad \underline{\int_R} f \leq \overline{\int_R} f.$$

Theorem 10.2.4:

Let $f: R \to \mathbb{R}$ be bounded. f is integrable iff for all $\varepsilon > 0$, there exists P such that
$$U(f,P) - L(f,P) < \varepsilon$$

Proof

$\Rightarrow$: Assume f is integrable. So:
$$\underline{\int_R} f = \overline{\int_R} f.$$
Let $\varepsilon > 0$ be given. We may find partitions P, Q,

such that.

$$\underline{\int_R} f - \frac{\varepsilon}{2} < L(f,P)$$

$$U(f,Q) < \overline{\int_R} f + \frac{\varepsilon}{2}$$

$\Rightarrow$.

$$U(f,Q) < L(f,P) + \varepsilon$$

Let S be a common refinement of P, Q. ($S \le P, Q$).

$$U(f,S) \le U(f,Q) < L(f,P) + \varepsilon \le L(f,S) + \varepsilon$$

$$\Rightarrow \quad U(f,S) - L(f,S) < \varepsilon$$

$\Leftarrow$: Assume there exists partition P such that

$$U(f,P) - L(f,P) < \varepsilon$$

for given $\varepsilon > 0$.

So:

$$0 \le \overline{\int_R} f - \underline{\int_R} f \le U(f,P) - L(f,P) < \varepsilon$$

$$\Rightarrow \overline{\int_R} f = \underline{\int_R} f.$$

$\square$

Module 3: Jordan Content and Lebesgue Measure.

Remark 10.3.1: Goal.

① Give a much stronger characterization of integrability

② To integrate over non-rectangles.

Definition 10.3.2:

① A set $A \subseteq \mathbb{R}^n$ has (Lebesgue) measure zero if for all $\varepsilon > 0$, there exists rectangles R_i $(i \in \mathbb{N})$ such that

$$A \subseteq \bigcup_{i=1}^{\infty} R_i$$

and

$$\sum_{i=1}^{\infty} V(R_i) < \varepsilon.$$

② A set $A \subseteq \mathbb{R}^n$ has (Jordan) content zero if for all $\varepsilon > 0$, there exists rectangles $R_1, R_2, \ldots, R_m$ such that

$$A \subseteq \bigcup_{i=1}^{m} R_i$$

and:

$$\sum_{i=1}^{m} V(R_i) < \varepsilon.$$

Proposition 10.3.3:

If $A \subseteq \mathbb{R}^n$ has content zero, then it has measure zero.

Proof:

Let $\varepsilon > 0$ be given. Suppose A has content zero. There exists rectangles $R_1, \ldots, R_m$ such that

$$A \subseteq R_1 \cup \cdots \cup R_m$$

and
$$\sum_{i=1}^{m} V(R_i) < \varepsilon.$$
For $i > m$, let $R_i \subseteq \mathbb{R}^n$ be any rectangle with volume zero (E.g take one side length to be $[a,a] = \{a\}$). So:
$$A \subseteq \bigcup_{i=1}^{\infty} R_i$$
and
$$\sum_{i=1}^{\infty} V(R_i) = \sum_{i=1}^{m} V(R_i) < \varepsilon. \qquad \square$$

Example 10.3.4 (measure zero but not content zero).

$A = \mathbb{Q} \subseteq \mathbb{R}^1$.

① Claim: $\mathbb{Q}$ has measure zero.

$$\mathbb{Q} = \{q_1, q_2, q_3, \ldots\}. \text{ (since } \mathbb{Q} \text{ is countable)}.$$

Let $\varepsilon > 0$ be given.

Let $R_i = [q_i - \frac{\varepsilon}{2^{i+2}}, q_i + \frac{\varepsilon}{2^{i+2}}]$.

Clearly $\mathbb{Q} \subseteq \bigcup_{i=1}^{\infty} R_i$. $\sum_{i=1}^{\infty} V(R_i) = \sum_{i=1}^{\infty} \frac{\varepsilon}{2^{i+1}} = \frac{\varepsilon}{2} < \varepsilon$.

This shows $\mathbb{Q}$ has measure zero.

② $\mathbb{Q}$ does not have Jordan content zero.

Take $\varepsilon = 1$. To cover $\mathbb{Q}$ by finitely many rectangles, $\mathbb{Q}$ has to be bounded.

Note: this proof can be used to show any countable set has Lebesgue measure zero.

Proposition 10.3.5:

If $A_1, A_2, \ldots, \subseteq \mathbb{R}^n$ have measure zero, then

$A = \bigcup_{i=1}^{\infty} A_i$ has measure zero.

Proof Sketch:

Let $\varepsilon > 0$ be given. $A_i \subseteq \bigcup_{j=1}^{\infty} R_{i,j}$, $\sum_{j=1}^{\infty} V(R_{i,j}) < \frac{\varepsilon}{2^i}$

$A \subseteq \bigcup_{i,j} R_{i,j}$. $\sum_{i,j} V(R_{i,j}) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} V(R_{i,j})$

$$< \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon.$$

Proposition 10.3.6:

If $A \subseteq \mathbb{R}^n$ is compact and has measure zero, then A has content zero.

Proof:

Let $\varepsilon > 0$ be given. By assignment, there exists <u>open</u> rectangles R_i such that

$$A \subseteq \bigcup_{i=1}^{\infty} R_i$$

and

$$\sum_{i=1}^{\infty} V(R_i) < \varepsilon$$

By compactness, $A \subseteq \bigcup_{i=1}^{m} R_i$ for some m.

Moreover, $$\sum_{i=1}^{m} V(R_i) \leq \sum_{i=1}^{\infty} V(R_i) < \varepsilon.$$

Module 4: Integrability and Measure.

Remark 10.4.1: Goal

Describe the relationship between integrability and measure zero.

Definition 10.4.2:

Let $A \subseteq \mathbb{R}^n$, $f: A \to \mathbb{R}$ be bounded. For $a \in A$ and $\delta > 0$, we define:

$$M(a, f, \delta) = \sup\{f(x) : x \in A, \|x-a\| < \delta\}$$

$$m(a, f, \delta) = \inf\{f(x) : x \in A, \|x-a\| < \delta\}.$$

$$\theta(f, a) = \lim_{\delta \to 0}(M(a, f, \delta) - m(a, f, \delta)).$$

$\theta(f, a)$ is called the oscillation of f at a.

Remark 10.4.3 [Piazza]

① $\theta(f, a)$ always exists

② f is continuous at a iff $\theta(f, a) = 0$.

Proposition 10.4.4.

Let $A \subseteq \mathbb{R}^n$ be closed. If $f: A \to \mathbb{R}$ is bounded, then $\forall \varepsilon > 0$, $\{x \in A : \theta(f, x) \geq \varepsilon\}$ is closed.

Proof:

Let $B = \{x \in A : \theta(f, x) \geq \varepsilon\}$. Take $x \in A \setminus B$.

So, $\theta(f, x) < \varepsilon$ and so there exists $\delta > 0$ with

$$M(x, f, \delta) - m(x, f, \delta) < \varepsilon$$

Consider $y \in B_{\delta/2}(x) \cap A$. Then, for $z \in A$ with $|y-z| < \frac{\delta}{2}$,

$$|z-x| \leq |z-y| + |y-x| < \delta$$

and so

$$m(x,f,\delta) \leq f(z) \leq M(x,f,\delta).$$

$\Rightarrow$ $M(y,f,\delta/2) - m(y,f,\delta/2) < \varepsilon$

$\Rightarrow$ $\theta(f,y) < \varepsilon$.

$\Rightarrow$ $B_{\delta/2}(x) \cap A \subseteq A \setminus B$.

$\Rightarrow$ $A \setminus B$ is relatively open in A.

$\Rightarrow$ B is relatively closed in A.

$\Rightarrow$ B is closed since A is closed. $\square$

Proposition 10.4.5.

Let $R \subseteq \mathbb{R}^n$ be rectangle, $f: R \to \mathbb{R}$ be bounded. Let $\varepsilon > 0$. If $\theta(f,x) < \varepsilon$ for all $x \in R$, then $\exists$ P such that $U(f,P) - L(f,P) < \varepsilon \cdot V(R)$.

Proof.

For all $x \in R$, we may find $\delta_x > 0$ such that

$$M(x,f,\delta_x) - m(x,f,\delta_x) < \varepsilon.$$

For all $x \in R$, let R_x be an open rectangle such that

$$x \in R_x \subseteq B_{\delta_x/2}(x).$$

Let $U_x = R_x \cap R$. We see that $R = \bigcup_{x \in R} U_x$ is a relatively open cover of R. Since R is compact,

there exists $x_1, \dots, x_m \in R$ such that $R = \mathcal{U}_{x_1} \cup \cdots \cup \mathcal{U}_{x_m}$.
Let P be a partition of R so fine that each subrectangle of P is contained in some $\overline{\mathcal{U}_{x_i}}$. Note

$$\overline{\mathcal{U}_{x_i}} = \overline{R_{x_i}} \cap R$$
$$\subseteq \overline{B_{\delta_{x_i}/2}(x_i)} \cap R$$
$$\subseteq B_{\delta_{x_i}}(x_i) \cap R.$$

So, for every $R_i \in P$

$$M_i - m_i < \varepsilon$$

$$\Rightarrow \quad U(f,P) - L(f,P) = \sum_{R_i \in P} (M_i - m_i) V(R_i)$$
$$< \sum_{R_i \in P} \varepsilon V(R_i) = \varepsilon V(R) \qquad \square.$$

Theorem 10.4.6.
Let $R \subseteq \mathbb{R}^n$ be rectangle, $f: R \to \mathbb{R}$ be bounded. Let $A = \{ x \in R : f \text{ is not continuous at } x \}$. Then f is integrable iff A has measure zero.

Proof: Appendix.

Week 10 Appendix

In the following, $R \subseteq \mathbb{R}^n$ will denote a rectangle and $f : R \to \mathbb{R}$ will be a bounded function. Also, we let

$$A = \{x \in R : f \text{ is not continuous at } x\}.$$

Theorem. If A has measure zero then f is integrable.

Proof. Suppose A has measure zero and let $\varepsilon > 0$ be given. Let $B = \{x \in R : o(f, x) \geq \varepsilon\}$, so that B is compact (R is bounded) from our first proposition. Since $B \subseteq A$ (remember that the oscillation function is zero at points of continuity), we have that B also has measure zero. Since B is compact, B also has content zero. In particular, we may find finitely many rectangles $U_1, \ldots, U_m$ whose interiors cover B (I am using an assignment problem here), such that $\sum v(U_i) < \varepsilon$.

Let X denote the set of subrectangles of R which are contained in at least one U_i. Let Y denote the set of subrectangles of R which are contained in $R \backslash B$. Now, since the U_i's cover B, we may find a partition $P = \{R_1, \ldots, R_k\}$ fine enough so that the elements of P are from either X or Y.

Since f is bounded, there exists M such that $|f(x)| \leq M$ for all $x \in R$. In particular, for every $R_i \in P$, $M_i - m_i \leq 2M$ (as in the definition of upper and lower sum). By the definition of X,

$$\sum_{R_i \in X} (M_i - m_i) v(R_i) \leq 2M \sum_{R_i \in X} v(R_i) \leq 2M \sum_{i=1}^{m} v(U_i) < 2M\varepsilon.$$

Now, if $R_i \in Y$ and $x \in R_i$, we have that $o(f, x) < \varepsilon$. By our second technical proposition (from class), we may find a partition $P_i = \{S_{i_1}, \ldots, S_{i_{\alpha(i)}}\}$ of R_i such that

$$\sum_{j=1}^{\alpha(i)} (M_j - m_j) v(S_{i_j}) < \varepsilon v(R_i).$$

By replacing each $R_i \in Y$ with $S_{i_1}, S_{i_2}, \ldots, S_{i_{\alpha(i)}}$ (and leaving the $R_i \in X$ alone), this creates a refinement $Q \leq P$. Finally,

$$\begin{aligned} U(f, Q) - L(f, Q) &= \sum_{R_i \in X} (M_i - m_i) v(R_i) + \sum_{R_i \in Y} \sum_{j=1}^{\alpha(i)} (M_j - m_j) v(S_{i_j}) \\ &< 2M\varepsilon + \sum_{R_i \in Y} \varepsilon v(R_i) \\ &\leq 2M\varepsilon + \varepsilon v(R). \end{aligned}$$

Since the final term can be made arbitrarily small, our function is integrable!

□

Theorem. If f is integrable then A has measure zero.

Proof. Suppose f is integrable and let $\varepsilon > 0$ be given. For every $n \in \mathbb{N}$, let

$$B_n = \{x \in R : o(f,x) \geq 1/n\}.$$

Since $A = B_1 \cup B_2 \cup \cdots$, it suffices to show that each B_n has measure zero. Fix $n \in \mathbb{N}$.

Since f is integrable, we may find a partition P of R such that $U(f,P) - L(f,P) < \varepsilon/n$. Let X denote the collection of rectangles in P which intersect B_n. In particular, the elements of X cover B_n and are rectangles! Now, if $R_i \in X$, then $M_i - m_i \geq 1/n$ by definition of the oscillation function. Then,

$$\begin{aligned} \frac{1}{n} \sum_{R_i \in X} v(R_i) &\leq \sum_{R_i \in X} (M_i - m_i) v(R_i) \\ &\leq \sum_{R_i \in P} (M_i - m_i) v(R_i) \\ &= U(f,P) - L(f,P) < \frac{\varepsilon}{n}. \end{aligned}$$

In particular, $\sum_{R_i \in X} v(R_i) < \varepsilon$ and so B_n has measure (content) zero.

$\square$

Week 11 Notes.

Module 0: Equivalent Notations.

	Old Notation.	New totation.
Partition.	Grid on R.	Partition of R.
Volume.	$\lvert R \rvert$.	$V(R)$.
Lower integral	$L\int_R f$.	$\underline{\int}_R f$.
Upper integral	$U\int_R f$.	$\overline{\int}_R f$.

Module 1: General Integrability.

Remark 11.1.1. Goal.
To be able to integrate over regions other than rectangles.

Definition 11.1.2:
Let $A \subseteq \mathbb{R}^n$ be bounded. Let R be a rectangle such that $A \subseteq R$. The characteristic function of A on R is:

$$\chi_A : R \to \mathbb{R}.$$

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A. \end{cases}$$

Definition 11.1.3:
Let $A \subseteq \mathbb{R}^n$ be bounded and R be a rectangle with $A \subseteq R$. We extend $f: R \to \mathbb{R}$ by setting $f(x) = 0$ for all $x \in R \setminus A$. We say $f: A \to \mathbb{R}$ is integrable iff. $f \cdot \chi_A : R \to \mathbb{R}$ is integrable, in which we define.

$$\int_A f = \int_R f \cdot \chi_A.$$

Assignment: This definition is independent of the choice of R containing A.

Remark 11.1.4:
When is $\chi_A : R \to \mathbb{R}$ integrable?

Theorem 11.1.5:

Let $A \subseteq \mathbb{R}^n$ be bounded, $A \subseteq R$ be a rectangle. The function $\chi_A : R \to \mathbb{R}$ is integrable iff $\partial(A)$ has measure zero.

Proof: Let $a \in R$.

① $a \in \text{Int}(A)$

Then, there exists an open ball $B_\delta(a) \subseteq A$. Since $\chi_A = 1$ on $B_\delta(a)$, χ_A is clearly continuous at a.

② $a \notin \overline{A}$

Then, $a \in \text{Int}(\mathbb{R}^n \setminus A)$ and so there exists an open ball $B_\delta(a) \subseteq \mathbb{R}^n \setminus A$. Since $\chi_A = 0$ on $B_\delta(a) \cap R$, χ_A is clearly continuous at a.

③ $a \in \overline{A} \setminus \text{Int}(A) = \partial(A)$.

Then, $a \in \overline{A}$ and $a \in \mathbb{R}^n \setminus \text{Int}(A) = \overline{\mathbb{R}^n \setminus A}$. In particular for all $\delta > 0$, there exists $x \in A$ and $y \in R \setminus A$ such that $\|x - a\|, \|y - a\| < \delta$. Thus,

$$\theta(\chi_A, a) \geq 1.$$

and so f is not continuous at a. The result follows from our big Theorem. $\square$

Definition 11.1.6:

① We call A a Jordan region iff $\partial(A)$ has measure zero. (iff χ_A is integrable on $A \subseteq R$).

② If A is a Jordan region with $A \subseteq R$, we define the volume of A by

$$Vol(A) = \int_R \chi_A = \int_A 1.$$

Proposition 11.1.7:

Let $A, B \subseteq \mathbb{R}^n$ be Jordan regions.

① $A \cup B$ is a Jordan region

② If. $A \cap B = \emptyset$ and $f: A \cup B \to \mathbb{R}$ is integrable, then.

$$\int_{A \cup B} f = \int_A f + \int_B f.$$

Proof sketch:

① $\partial(A \cup B) = \overline{(A \cup B)} \setminus Int(A \cup B).$

$$\subseteq (\overline{A} \cup \overline{B}) \setminus (Int(A) \cup Int(B)).$$

$$\subseteq (\overline{A} \setminus Int(A)) \cup (\overline{B} \setminus Int(B)).$$

$$= \partial(A) \cup \partial(B)$$

② Let $A \cup B \subseteq R$.

$$\int_{A \cup B} f = \int_R f \chi_{A \cup B} = \int_R f (\chi_A + \chi_B)$$

$$= \int_R f \chi_A + \int_R f \chi_B$$

$$= \int_A f + \int_B f.$$

Module 2 : Fubini's Theorem.

Remark 11.2.1:

Can we use multiple single-variable integrals to compute integrals of multivariable functions? Would the order matter?

i.e. $\int_c^d \int_a^b f(x,y)\, dx\, dy = \int_a^b \int_c^d f(x,y)\, dy\, dx$??

Remark 11.2.2:

Let $B \subseteq \mathbb{R}^2$ be Jordan region.

$f: B \to \mathbb{R}$ be integrable.

$$\int_B f(v)\, dv \equiv \iint_B f(x,y)\, dA.$$

$B \subseteq \mathbb{R}^3$ be Jordan region.

$f: B \to \mathbb{R}$ be integrable.

$$\int_B f(v)\, dv \equiv \iiint_B f(x,y,z)\, dV.$$

Lemma 11.2.3.

Let $R = [a,b] \times [c,d] \subseteq \mathbb{R}^2$, $f: R \to \mathbb{R}$ be bounded.

If $f(x,\cdot): [c,d] \to \mathbb{R}$ given by $f(x,\cdot)(y) = f(x,y)$ is integrable for all $x \in [a,b]$, then

$$L\iint_R f(x,y)\, dA \leq L\int_a^b \left(\int_c^d f(x,y)\, dy\right) dx$$

$$\leq U\int_a^b \left(\int_c^d f(x,y)\, dy\right) dx$$

$$\leq U\iint_R f(x,y)\, dA.$$

Proof:

The middle inequality is trivial. We will prove the last inequality and leave the first for discussion.

Let $\varepsilon > 0$ be given. Choose a partition G on R such that $U(f,G) - \varepsilon \leq U\iint_R f(x,y)\,dA$.

Say

$$G = \{R_{ij} : 1 \leq i \leq k,\ 1 \leq j \leq l\}.$$

$$R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j].$$

where $x_0 = a$, $x_k = b$, $y_0 = c$, $y_l = b$.

Set $M_{ij} = \sup\{f(v) : v \in R_{ij}\}$.

$$\begin{aligned}
&U\int_a^b \left(\int_c^d f(x,y)\,dy\right)dx \\
&= \sum_{i=1}^k U\int_{x_{i-1}}^{x_i} \left(\sum_{j=1}^l \int_{y_{j-1}}^{y_j} f(x,y)\,dy\right)dx \\
&\leq \sum_{i=1}^k \sum_{j=1}^l U\int_{x_{i-1}}^{x_i} \left(\int_{y_{j-1}}^{y_j} f(x,y)\,dy\right)dx
\end{aligned}$$

(Here we use $U\int_a^b (f+g)(x)\,dx \leq U\int_a^b f(x)\,dx + U\int_a^b g(x)\,dx$)

$$\begin{aligned}
&\leq \sum_{i=1}^k \sum_{j=1}^l \int_{x_{i-1}}^{x_i}\left(\int_{y_{j-1}}^{y_j} M_{ij}\,dy\right)dx. \\
&= \sum_{i=1}^k \sum_{j=1}^l M_{ij}\,(x_i - x_{i-1})(y_j - y_{j-1}). \\
&= \sum_{R_{ij}} M_{ij}\,|R_{ij}|. \\
&= U(f,G) \\
&\leq U\iint_R f(x,y)\,dA + \varepsilon
\end{aligned}$$

$\square$.

Theorem 11.2.4 [Fubini's Theorem].

Let $R = [a,b] \times [c,d] \subseteq \mathbb{R}^2$.

$f : R \to \mathbb{R}$ be integrable.

If $f(x,\cdot)$ and $f(\cdot,y)$ are integrable over $[c,d]$ and $[a,b]$ respectively, for all $x \in [a,b]$ and $y \in [c,d]$, then.

$$\iint_R f(x,y)\, dA = \int_a^b \int_c^d f(x,y)\, dy\, dx$$
$$= \int_c^d \int_a^b f(x,y)\, dx\, dy.$$

Proof:

Since f is integrable

$$L\iint_R f(x,y)\, dA = U\iint_R f(x,y)\, dA.$$

By the Lemma, this implies

$$L\int_a^b \int_c^d f(x,y)\, dy\, dx = U\int_a^b \int_c^d f(x,y)\, dy\, dx$$

So:

$$\iint_R f(x,y)\, dA = \int_a^b \left(\int_c^d f(x,y)\, dy\right) dx.$$

Rearrange the roles of x, y in the lemma, we proves the theorem. $\square$

Remark 11.2.5:

The conditions of Fubini's Theorem is met when f is continuous.

Remark 11.2.6:

We call

$$\int_{a_n}^{b_n}\int_{a_{n-1}}^{b_{n-1}}\cdots\int_{a_1}^{b_1} f(x_1, x_2, \dots, x_n)\, dx_1\, dx_2 \cdots dx_n$$

iterated integrals.

Example 11-2.7:

$R = [1,2] \times [0,\pi]$.

$$\iint_R y \sin(xy)\, dA$$

[Note: f, $f(x,\cdot)$, $f(\cdot,y)$ are all continuous on closed Jordan Regions $\Rightarrow$ integrable].

Fubini's Theorem $\displaystyle\int_0^{\pi}\int_1^{2} y\sin(xy)\, dx\, dy$

$$= \int_0^{\pi} \left(-\cos(xy)\right)_{x=1}^{x=2} dy$$

$$= \int_0^{\pi} -\cos 2y + \cos y \, dy.$$

$$= \left[-\frac{1}{2}\sin 2y + \sin y\right]_0^{\pi}$$

$$= 0.$$

Module 3: Iterated Integrals.

Remark 11.3.1: Goal

Generalizing Fubini's Theorem.

Theorem 11.3.2:

$R = [a_1, b_1] \times \cdots \times [a_n, b_n] \subseteq \mathbb{R}^n$.

$R_n = [a_1, b_1] \times \cdots \times [a_{n-1}, b_{n-1}]$

If $f(x, \cdot)$ is integrable for all $x \in R_n$, then $\int_{a_n}^{b_n} f(x,t)\,dt$ is integrable on R_n and

$$\int_R f(v)\,dv = \int_{R_n} \int_{a_n}^{b_n} f(x,t)\,dt\,dx.$$

Remark 11.3.3:

If $f: R \to \mathbb{R}$ is continuous, then

$$\int_R f(v)\,dv = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} f(x_1, \dots, x_n)\,dx_n \cdots dx_1$$

Remark 11.3.4:

We will use iterated integrals to integrate over "nice" regions which are not rectangles.

For simplicity, we shall work in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Definition 11.3.5:

① We say $A \subseteq \mathbb{R}^2$ is type 1 if

$$A = \{(x,y) : x \in [a,b],\ \varphi(x) \le y \le \psi(x)\}.$$

for some continuous $\varphi, \psi : [a,b] \to \mathbb{R}$.

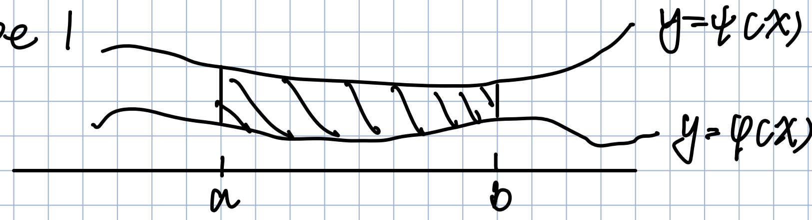
② We say $A \subseteq \mathbb{R}^2$ is type 2 if

$$A = \{(x,y) : y \in [a,b],\ \varphi(y) \leq x \leq \psi(y)\}$$

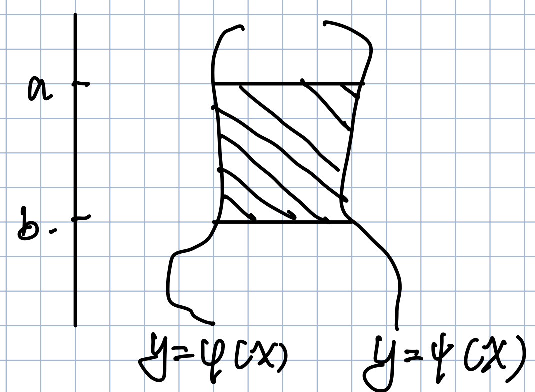
for some continuous $\varphi, \psi : [a,b] \to \mathbb{R}$.

Example 11.3.6: In $\mathbb{R}^2$:

Type 1



Type 2:



Definition 11.3.6: $A \subseteq \mathbb{R}^3$.

① Type 1.

$$A = \{(x,y,z) : (x,y) \in H,\ \varphi(x,y) \leq z \leq \psi(x,y)\}.$$

② Type 2:

$$A = \{(x,y,z) : (x,z) \in H,\ \varphi(x,z) \leq y \leq \psi(x,y)\}.$$

③ Type 3:

$$A = \{(x,y,z) : (y,z) \in H,\ \varphi(y,z) \leq x \leq \psi(y,z)\}.$$

where $H \subseteq \mathbb{R}^2$ is a closed Jordan region and $\varphi, \psi : H \to \mathbb{R}$ are continuous.

Remark 11.3.7:
Regions of type 1, 2, or 3 are Jordan regions.

Theorem 11.3.8:
$A \subseteq \mathbb{R}^2$, $f: A \to \mathbb{R}$ is continuous.

① If A is type 1, so that.
$$A = \{(x,y) : x \in [a,b], \varphi(x) \le y \le \psi(x)\}.$$
for some continuous $\varphi, \psi : [a,b] \to \mathbb{R}$.
then
$$\int_A f(v)\,dV = \int_a^b \int_{\varphi(x)}^{\psi(x)} f(x,y)\,dy\,dx.$$

② If A is type 2, so that
$$A = \{(x,y) : y \in [a,b], \varphi(y) \le x \le \psi(y)\}.$$
then
$$\int_A f(v)\,dv = \int_a^b \int_{\varphi(y)}^{\psi(y)} f(x,y)\,dx\,dy.$$

Proof (of ①).
Let $R = [a,b] \times [c,d]$ be a rectangle containing A.
Extend f to R by setting $f = 0$ on $R \setminus A$. By Fubini's Theorem,
$$\int_A f(v)\,dV = \int_R f(v)\,dV$$
$$= \int_a^b \int_c^d f(x,y)\,dy\,dx.$$

However, $f(x,y) = 0$ if it is not the case that
$$\varphi(x) \le y \le \psi(x).$$
So:
$$\int_A f(v)\,dV = \int_a^b \int_{\varphi(x)}^{\psi(x)} f(x,y)\,dy\,dx \qquad \square$$

Theorem 11.3.9:

$A \subseteq \mathbb{R}^3$, $f: A \to \mathbb{R}$ is continuous

If A is type 1, then.

$$\int_A f(v)\, dv = \int_H \int_{\varphi(x,y)}^{\psi(x,y)} f(u, z)\, dz\, du$$

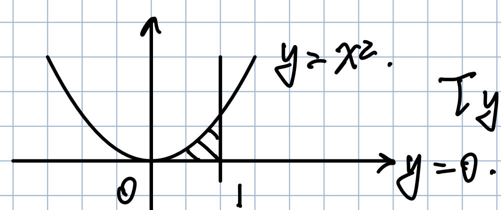
Tybe 2, Type 3 are similar.

Module 4: Examples.

Example 11.4.1:

Let $D \subseteq \mathbb{R}^2$ be the region be the region bounded by $y=0$, $y=x^2$, $x=1$. Compute

$$\iint_D x \cos y \, dA.$$



Type 1.

$$\iint_D x\cos y \, dA = \int_0^1 \int_0^{x^2} x\cos y \, dy \, dx.$$

$$= \int_0^1 \left[x \sin y \right]_{y=0}^{y=x^2} dx.$$

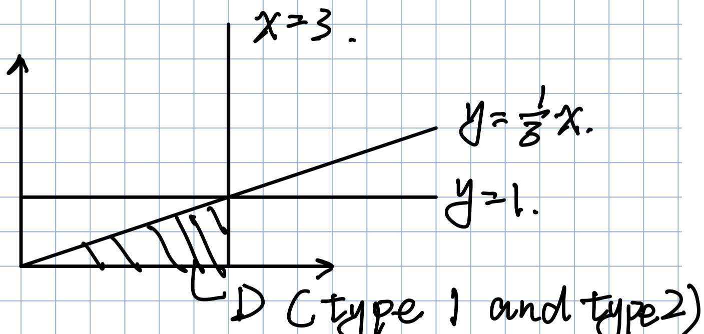
$$= \int_0^1 x \sin x^2 \, dx.$$

$$= \left[-\frac{1}{2}\cos x^2 \right]_0^1.$$

$$= -\frac{1}{2}\cos(1) + \frac{1}{2}$$

Example 11.4.2:

$$\int_0^1 \int_{3y}^3 e^{x^2} dx\, dy$$



$$\int_0^1 \int_{3y}^3 e^{x^2}\,dx\,dy = \iint_D e^{x^2}\,dA.$$

$$= \int_0^3 \int_0^{\frac{1}{3}x} e^{x^2}\,dy\,dx$$

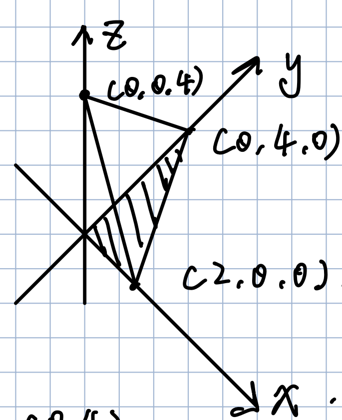
$$= \int_0^3 \left[e^{x^2} y\right]_{y=0}^{y=\frac{1}{3}x} dx$$

$$= \int_0^3 \frac{1}{3} x e^{x^2}\,dx$$

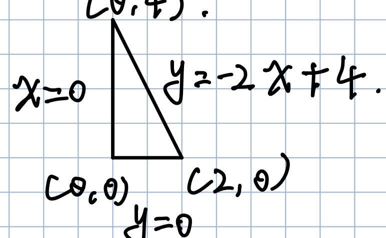
$$= \left[\frac{1}{6} e^{x^2}\right]_0^3 = \frac{1}{6}e^9 - \frac{1}{6}$$

Example 11.4.3:

Find the volume of the tetrahedron T enclosed by $x=0$, $y=0$, $z=0$ and $2x+y+z=4$.



The base:



$$T = \{(x,y,z) : 0 \le x \le 2,\ 0 \le y \le -2x+4,\ 0 \le z \le 4-2x-y\}.$$

$H = \{(x,y) : 0 \le x \le 2,\ 0 \le y \le -2x+4\}.$

$\hookrightarrow$ Type I.

$T = \{(x,y,z) : (x,y) \in H,\ 0 \le z \le 4-2x-y\}.$

$$\iiint_T 1 \cdot dV$$

$$= \int_H \underbrace{\int_0^{4-2x-y} 1 \cdot dz}_{\text{continuous}}\, dA.$$

$$= \int_0^2 \int_0^{-2x+4} \int_0^{4-2x-y} 1 \cdot dz\, dy\, dx.$$

$$= \int_0^2 \int_0^{-2x+4} 4-2x-y \; dy\, dx$$

$$= \int_0^2 \left[(4-2x)y - \frac{1}{2}y^2\right]_{y=0}^{y=4-2x} dx$$

$$= \int_0^2 \frac{1}{2}(4-2x)^2\, dx = \frac{16}{3}$$

Module 5: Change of Variables.

Remark 11.5.1:
Recall $f: A \to \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$, $a \in A$. The Jacobian of f at a is:

$$Jf(a) = \det(Df(a)).$$

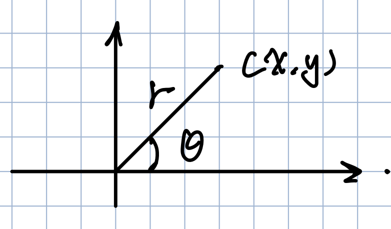
Theorem 11.5.2:
Let $U \subseteq \mathbb{R}^n$ be open, $A \subseteq U$ be closed Jordan region. Let $f: A \to \mathbb{R}$ be continuous and let $\varphi \in C^1(U, \mathbb{R}^n)$. Suppose $\exists\, B \subseteq A$.

1. $\mathrm{Vol}(B) = 0$.
2. φ is injective on $A \setminus B$.
3. $J\varphi(a) \neq 0$, $\forall\, a \in A \setminus B$.

and suppose $f: \varphi(A) \to \mathbb{R}$ is continuous. Then $\varphi(A)$ is a Jordan region, f is injective on $\varphi(A)$ and

$$\int_{\varphi(A)} f(x)\,dx = \int_A f(\varphi(x))\,|J\varphi(x)|\,dx.$$

Theorem 11.5.3: Polar Coordinates.



$r = \sqrt{x^2 + y^2}$
$x = r\cos\theta$
$y = r\sin\theta$
$0 \leq \theta \leq 2\pi$.

We call (r, θ) the polar coordinates of $(x, y) \in \mathbb{R}^2$.
Consider $\varphi \in C^1(\mathbb{R}^2, \mathbb{R}^2)$ given by

$$\varphi(r,\theta) = (r\cos\theta, r\sin\theta).$$

Note: φ is injective on $\mathbb{R}^2 \setminus \underbrace{\{(0,\theta): 0 \le \theta \le 2\pi\}}_{\text{Volume zero.}}$.

$$|J\varphi(r,\theta)| = \left|\det\begin{bmatrix}\cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta\end{bmatrix}\right| = |r| = r.$$

$$\iint_{\varphi(D)} f(x,y)\,dA = \iint_D f(r\cos\theta, r\sin\theta)\, \underset{\sim}{r}\, dA$$

Example 11.5.4:

$$\iint_D \cos(x^2+y^2)\,dA.$$

D is the region bounded by $x^2+y^2=9$ and above the x-axis.

$$D = \{(r,\theta): 0 \le r \le 3,\ 0 \le \theta \le \pi\}.$$

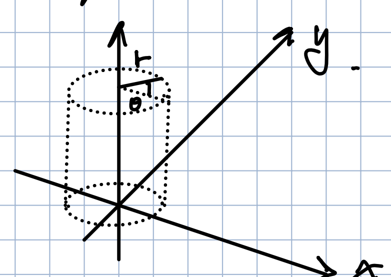
$$\iint_D \cos(x^2+y^2)\,dA = \iint_D \cos(r^2)\, r\, dA.$$

$$= \int_0^{\pi}\int_0^3 \cos(r^2)\, r\, dr\, d\theta$$

$$= \int_0^{\pi} \left[\frac{1}{2}\sin(r^2)\right]_0^3 dr$$

$$= \int_0^{\pi} \frac{1}{2}\sin 9\,\theta = \frac{\pi}{2}\sin 9.$$

Theorem 11.5.5. Cylinderical Coordinates.



We call (r, θ, z) the cylinderical coordinates of (x, y, z).

$$\varphi(r, \theta, z) = (r\cos\theta, r\sin\theta, z).$$

$$|J\varphi(r,\theta,z)| = \left| \det \begin{bmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \right|$$

$$= r.$$

$$\iiint_{\varphi(A)} f(x,y,z)\, dV = \iiint_{A} f(r\cos\theta, r\sin\theta, z)\, r\, dV$$

Example 11.5-6.

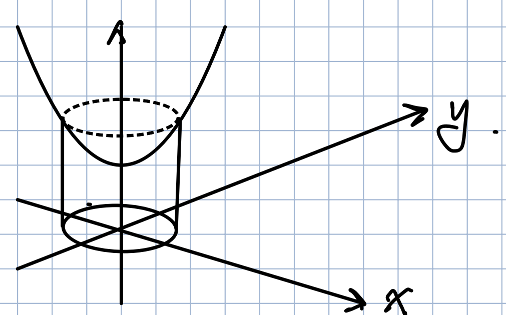
$$\iiint_A e^z \, dV$$

where A is enclosed by

① The paraboloid $z = 1 + x^2 + y^2$.

② The cylinder $x^2 + y^2 = 5$.

③ The xy-plane.



$$A = \{(r, \theta, z) : 0 \le r \le \sqrt{5},\ 0 \le \theta \le 2\pi,\ 0 \le z \le 1 + r^2\}$$

$$\iiint_A e^z \, dV = \int_0^{\sqrt{5}} \int_0^{2\pi} \int_0^{1+r^2} e^z r \, dz\, d\theta\, dr$$

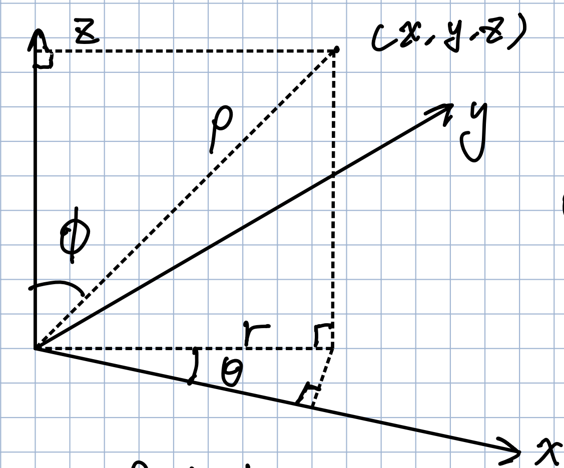
$$= \int_0^{\sqrt{5}} \int_0^{2\pi} r e^{1+r^2} - r \, d\theta\, dr.$$

$$= 2\pi \int_0^{\sqrt{5}} r e^{1+r^2} - r \, dr.$$

$$= 2\pi \left[\frac{1}{2} e^{1+r^2} - \frac{1}{2} r^2 \right]_0^{\sqrt{5}}$$

$$= \pi (e^6 - 5 - e).$$

Theorem 11.5.7: Spherical Coordinates.



$0 \leq \theta \leq 2\pi$

$0 \leq \phi \leq \pi.$

$x = r\cos\theta \qquad r = \rho\sin\phi$

$y = r\sin\theta \qquad z = \rho\cos\phi.$

$x = \rho\sin\phi\cos\theta, \; y = \sin\theta\sin\phi.$

$z = \rho\cos\phi, \quad x^2 + y^2 + z^2 = \rho^2.$

Consider $\varphi(\rho, \theta, \phi) = (\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi)$

$$|J\varphi(\rho, \theta, \phi)| = \left| \det \begin{bmatrix} \sin\phi\cos\theta & -\rho\sin\phi\sin\theta & \rho\cos\phi\cos\theta \\ \sin\phi\sin\theta & \rho\sin\phi\cos\theta & \rho\cos\phi\sin\theta \\ \cos\phi & 0 & -\rho\sin\phi \end{bmatrix} \right|$$

$$= \rho^2 \sin\phi.$$

$$\iiint_{\varphi(A)} f(x, y, z)\, dV = \iiint_A f(\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi)\rho^2\sin\phi\, dV$$

Example 11.5.8.

Find the volume of the sphere $x^2+y^2+z^2=a^2$.

$S=\{(\rho,\theta,\phi): 0\le\rho\le a,\ 0\le\theta\le 2\pi,\ 0\le\phi\le\pi\}$.

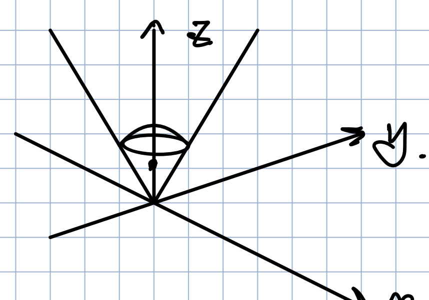
$$\begin{aligned}
Vol(S) &= \int_S 1\,dV \\
&= \iiint_S 1\,dx\,dy\,dz \\
&= \int_0^a\int_0^{2\pi}\int_0^{\pi} \rho^2\sin\phi\,d\phi\,d\theta\,dr. \\
&= \int_0^a\int_0^{2\pi} \left[-\rho^2\cos\phi\right]_0^{\pi}\,d\theta\,d\rho \\
&= \int_0^a\int_0^{2\pi} 2\rho^2\,d\theta\,d\rho \\
&= 2\pi\int_0^a 2\rho^2\,d\rho \\
&= 2\pi\cdot\frac{2}{3}\rho^3\Big|_0^a = \frac{4\pi}{3}a^3.
\end{aligned}$$

Example 11.5.9.

Find the volume of the solid which

① lies above the cone $z=\sqrt{x^2+y^2}$.

② below the sphere $x^2+y^2+z^2=z$.



$$x^2+y^2+z^2=z \iff x^2+y^2+(z-\frac{1}{2})^2=\frac{1}{4}$$

Cone: $\rho\cos\phi = \sqrt{\rho^2\sin^2\phi\cos^2\theta + \rho^2\sin^2\phi\sin^2\theta}$.

$= \rho\sin\phi$.

$C = \{(\rho,\theta,\phi) : \rho = 0 \text{ or } \phi = \frac{\pi}{4}\}$.

Sphere: $\rho^2 = \rho\cos\phi$

$S = \{(\rho,\theta,\phi) : \rho = 0 \text{ or } \rho = \cos\phi\}$.

Let D be the above solid.

$$\iiint_D 1\, dV = \iiint_D \rho^2 \sin\phi \, d\rho.$$

$$= \int_0^{2\pi}\int_0^{\pi/4}\int_0^{\cos\phi} \rho^2\sin\phi\, d\rho\, d\phi\, d\theta.$$

$$= \frac{\pi}{8}.$$