Class Notes

数分 III

Math 247

Calculus III

Advanced Level

Spring 2020

By 李泽天

1.1 Normed Vector Spaces

<u>Intro to Real Analysis</u> : Analysis is the study of approximation of mathematical objects

<u>Idea</u> : [Normed Vector Space] : A NVS is a vector space where we can measure the distance between vectors

<u>Def'n</u> : Let V be a real vector space.

A **norm** on V is a function : $\|\cdot\| : V \to \mathbb{R}$

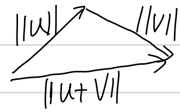
such that ① $\|v\| \geq 0$ for all $v \in V$

② $\|v\| = 0$ iff $v = 0$

③ For all $\alpha \in \mathbb{R}$, $v \in V$: $\|\alpha v\| = \|\alpha\| \cdot \|v\|$

④ Triangle Inequality : For all $u, v \in V$: $\|u+v\| \leq \|u\| + \|v\|$

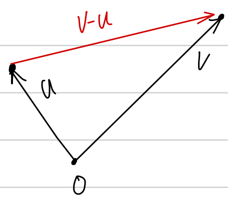
<u>Aside</u> :



Let $\|\cdot\|$ be a norm on V, we call the pair $(V, \|\cdot\|)$ a <u>normed vector space</u>

<u>Convention</u> : If $\|\cdot\|$ is understood, we write V instead of $(V, \|\cdot\|)$

<u>idea</u> : (in $\mathbb{R}^2$)



$\|v-u\|$ = distance between v, u

$\|v\|$ = "length" of v

$\|v-0\|$ = distance between $v, 0$

1.2 Examples

Ex: $(\mathbb{R}, |\cdot|)$
↳ absolute value

Ex: $(\mathbb{R}^n, \|\cdot\|_2)$: $\|(x_1, x_2, \cdots x_n)\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ Euclidean Norm
two norm

Ex: $(\mathbb{R}^n, \|\cdot\|)$: $p \geq 1$, $p \in \mathbb{R}$: $\left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{\frac{1}{p}}$ (See this in A1)
p-norm
Little home work: why we can't use $p<1$?

Ex: $(\mathbb{R}^n, \|\cdot\|_\infty)$: $\|(x_1, x_2, \cdots, x_n)\|_\infty = \sup\{|x_i| : i = 1, 2, \cdots, n\} = \max\{|x_i| : i = 1, 2, \cdots, n\}$
Sup norm / infinity norm

Ex: $\mathbb{R}^N := \{(x_i)_{i=1}^\infty : x_i \in \mathbb{R}\}$

$p \geq 1$ (real number) : $\|(x_i)_{i=1}^\infty\|_p = \left(\sum_{i=1}^\infty |x_i|^p\right)^{\frac{1}{p}}$ $(\stackrel{?}{=} \infty)$

$\ell^p := \{(x_i) \in \mathbb{R}^N : \|(x_i)\|_p < \infty\}$

↳ subspace of $\mathbb{R}^N$ (A1: $(\ell^p, \|\cdot\|_p)$ is a NVS)

Ex: $(x_i) \in \mathbb{R}^N$, $\|(x_i)\|_\infty = \sup\{|x_i| : i \in \mathbb{N}\}$ $(\stackrel{?}{=} \infty)$

$\ell^\infty = \{(x_i) \in \mathbb{R}^N : \|(x_i)\|_\infty < \infty\}$

↳ subspace of $\mathbb{R}^N$ (A1: $\ell^\infty, \|\cdot\|_\infty$ is NVS) Sub / infinity norm

Ex: $a<b$, $C([a,b]) = \{f : [a,b] \to \mathbb{R} \text{ continuous}\}$: $\|f\|_\infty = \sup\{|f(x)| : x \in [a,b]\}$

$\stackrel{E.V.T}{=} \max\{|f(x)| : x \in [a,b]\}$

$(C([a,b]), \|\cdot\|_\infty)$ is NVS
↓
uniform norm

1.3 Convergence

Def'n : Let V be NVS

A sequence in V is a right-infinite ordered list $(V_1, V_2, \cdots)$ where each $V_i \in V$

We denote this sequence by $(V_i)_{i=1}^{\infty}$ or (V_i), we also write $V_i \subseteq V$ to mean each $V_i \in V$

Def'n : Let V be NVS, $(a_n) \subseteq V$, $v \in V$

We say (a_n) Converges to v, written $a_n \to v$, if for all $\varepsilon > 0$, there exist $N \in \mathbb{N}$, such that if $n \geq N$, then $\|a_n - v\| < \varepsilon$

We call v is the limit of (a_n)

If (a_n) does not Converge to any $v \in V$, we say (a_n) diverges (in V)

Ex : $V = \ell^{\infty} = \{(x_i) \in \mathbb{R}^{\mathbb{N}} : \sup_{i \in \mathbb{N}} \{|x_i|\} < \infty\}$

$(a_n) \subseteq V$ where $a_n = (1, \frac{1}{2}, \frac{1}{3}, \cdots\cdots, \frac{1}{n}, 0, 0, 0, \cdots)$

☆ Claim : $a_n \to (1, \frac{1}{2}, \frac{1}{3}, \cdots)$

Let $\varepsilon > 0$ be given, choose $N = \frac{1}{\varepsilon}$ and suppose $n \geq N$

Then,

$$\begin{aligned}
&\|a_n - (1, \tfrac{1}{2}, \tfrac{1}{3}, \cdots)\|_{\infty} \\
&= \|(0, 0, \cdots, 0, -\tfrac{1}{n+1}, -\tfrac{1}{n+2}, \cdots)\|_{\infty} \\
&= \sup\{0, \tfrac{1}{n+1}, \tfrac{1}{n+2}, \cdots\} \\
&= \tfrac{1}{n+1} \\
&< \tfrac{1}{n} \leq \tfrac{1}{N} = \varepsilon
\end{aligned}$$

This proves the claim

Remark : By replacing N with $\lceil N \rceil$, we may replace $N \in \mathbb{N}$ with $N \in (0, \infty)$ in the def'n of Convergence

↳ ceiling function

1.4 More Convergence

Ex: $V = \ell^\infty$. $a_n \subseteq \ell^\infty$. $a_n = (1, 2, \cdots, n, 0, 0, \ldots)$

Claim: (a_n) diverges in ℓ^∞

For contradiction, suppose there exist $V = (V_1, V_2, \cdots)$ in ℓ^∞ s.t. $a_n \to V$

Consider $\varepsilon = 1$, there exist $N \in \mathbb{N}$ s.t if $n \geq N$ then $\|a_n - V\|_\infty < 1$

Since $V \in \ell^\infty$, there exist $M \in \mathbb{N}$ s.t. $|V_i| \leq M$ for all $i \in \mathbb{N}$, assume that $M > N$

observe that, $1 > \|a_{M+1} - V\|_\infty$

$$= \|(1, 2, \cdots, M+1, 0, \cdots) - (V_1, V_2, \cdots)\|_\infty$$
$$\geq |M+1 - V_{M+1}|$$
$$\geq |M+1| - |V_{M+1}|$$
$$\geq |M+1| + M$$
$$= 1$$

$1 > 1$ Contradiction

Ex: in $(C_{([0,1])}, \|\cdot\|_\infty)$: $(f_n) \subseteq C_{([0,1])}$, $f_n(x) = (x - \frac{1}{n})^2$

Claim: $f_n \to f$, $f(x) = x^2$

Let $\varepsilon > 0$ be given, observe that for $x \in [0,1]$, $|f_n(x) - f(x)| = |(x - \frac{1}{n})^2 - x^2| = |-\frac{2}{n}x + \frac{1}{n^2}|$

$$\leq \frac{2}{n}|x| + \frac{1}{n^2}$$
$$\leq \frac{2}{n} + \frac{1}{n^2} \to 0$$

Thus, $\exists N \in \mathbb{N}$ s.t. if $n \geq N$, then $\frac{2}{n} + \frac{1}{n^2} < \varepsilon$

so for $n \geq N$, $|f_n(x) - f(x)| < \varepsilon$

By the def'n of sup, $\|f_n - f\|_\infty < \varepsilon$

Prop: Let V be NVS, $a_n, b_n \subseteq V$, suppose $a_n \to v \in V$ and $b_n \to w \in V$, then ① $a_n + b_n \to v + w$

② $\lambda a_n \to \lambda v$ $(\lambda \in \mathbb{R})$

Proof: homework practice

2.1 Cauchy Sequence

Problem: The def'n of Converge requires the limit of the sequence

Goal: Find a new and equivalent notation of convergence which does not involve the limit of the sequence

Prop: Let V be Normed vector Space, $(a_n) \subseteq V$, $a_n \to a \in V$.

For all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ s.t. for all $n, m \geq N$, $\|a_n - a_m\| < \varepsilon$

Proof: Let $\varepsilon > 0$ be given, Since $a_n \to a$, there exist $N \in \mathbb{N}$ s.t if $n \geq N$, then $\|a_n - a\| < \frac{\varepsilon}{2}$

Then, if $n, m \geq N$, then $\|a_n - a_m\| = \|a_n - a + a - a_m\| \leq \|a_n - a\| + \|a - a_m\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ Converse true

Def'n: Let V be a Normed Vector space, $(a_n) \subseteq V$, we say (a_n) is Cauchy

if $\forall \varepsilon > 0$ $\exists N \in \mathbb{N}$ s.t. whenever $n, m \geq N$, $\|a_n - a_m\| < \varepsilon$

Remark: Convergent $\Rightarrow$ Cauchy Sequence

ex. $V = C_{\infty} := \{(a_n) \in \ell^{\infty} : \exists N \in \mathbb{N}, \forall n \geq N, a_n = 0\}$

equip V with $\|\cdot\|_{\infty}$

Consider $(a_n) \subseteq V$ given by $a_n = (1, \frac{1}{2}, \ldots, \frac{1}{n}, 0, 0, \ldots)$

we showed that $a_n \to a \in \ell^{\infty}$ where $a = (1, \frac{1}{2}, \frac{1}{3}, \ldots) \notin C_{\infty}$

$\therefore (a_n) \subseteq C_{\infty}$ diverges

Claim: (a_n) is Cauchy. (a_n) convergent in ℓ^{∞} $\Rightarrow$ (a_n) Cauchy in ℓ^{∞}

$\Rightarrow$ a_n Cauchy in C_{∞}

2.2 Completeness

Def'n Let V be normed vector spaces, $(a_n) \subseteq V$, we say (a_n) is bounded if $\exists N \in \mathbb{N}$ s.t. $\|a_n\| < N$

Ex: $(-1)^n \subseteq \mathbb{R}$, it's bounded, but divergent

Prop: Let V be normed vector spaces, if $a_n \subseteq V$ is Cauchy, then a_n is bounded   Converse not true

Proof: Suppose a_n is Cauchy, Consider $\varepsilon = 1$ so that $\exists N \in \mathbb{N}$ s.t. $\|a_n - a_m\| < 1$ for all $n, m \geq N$

For $n \geq N$, $\|a_n - a_N\| < 1 \implies \|a_n\| - \|a_N\| \leq \|a_n - a_N\| < 1$

$\implies \|a_n\| < 1 + \|a_N\|$

Let $M = \max\{\|a_1\|, \cdots, \|a_{N-1}\|, 1 + \|a_N\|\}$, so that $\|a_n\| \leq M$ for all $n \in \mathbb{N}$ $\square$

Idea: Convergent $\neq$ Cauchy   But Some time it is ----

Def'n: Let V be normed vector spaces, we say $A \subseteq V$ is complete if every Cauchy sequence $(a_n) \subseteq A$ converges in A

If V is complete itself (i.e. $A = V$) we call V is a Banach space

ex) $\mathbb{R}$ complete
(Banach space)

ex) $\mathbb{R}^n$ is Banach space

ex) ℓ^∞ is Banach space

ex) C_{00} Not Banach space

ex) $(0,1) \subseteq \mathbb{R}$, $(\frac{1}{n+1}) \subseteq (0,1)$, $\frac{1}{n+1} \to 0 \notin (0,1)$. $(\frac{1}{n+1})$ Convergent in $\mathbb{R}$

$\implies (\frac{1}{n+1})$ Cauchy

Since $0 \notin (0,1)$, so $(0,1)$ is not Complete

2.3 Topology 1

Roughly speaking, topology is the study of subsets of a set X which afford X meaningful analytic/geometric properties

Big idea: Given a normed vector space V, we want to investigate the way convergence/limits of sequences behave in subsets of V

Def'n: Let V be normed vector spaces
We say $C \subseteq V$ is closed (in V) if whenever $(x_n) \subseteq C$ s.t. $x_n \to x \in V$, then $x \in C$

idea: C is closed iff C is "closed" under taking limits

examples: ① $\emptyset, V \subseteq V$ closed

② Let V be normed vector spaces, $x \in V \Rightarrow \{x\}$ is closed why? (x): $x \to x \in \{x\}$

③ $[0,1) \subseteq \mathbb{R}$, $(1-\frac{1}{n}) \subseteq [0,1)$, $1-\frac{1}{n} \to 1 \notin [0,1)$
$\therefore [0,1)$ is not closed

④ $C_{00} \subseteq \ell^\infty$, $a_n = (1, \frac{1}{2}, \cdots, \frac{1}{n}, 0, \cdots) \in C_{00}$, $a_n \to (1, \frac{1}{2}, \frac{1}{3}, \cdots) \notin C_{00}$
$\therefore C_{00}$ is not closed in ℓ^∞

⑤ $C_{00} \subseteq C_{00}$, C_{00} is closed in C_{00} (depends on the space)

⑥ $A = \{f \in C_{([0,1])} : f(\frac{1}{2}) = 0\} \subseteq C_{([0,1])}$

Claim: A is closed in $C_{([0,1])}$.

Proof: Let $(f_n) \subseteq A$ s.t. $f_n \to f \in C_{([0,1])}$, we will show $f \in A$

Let $\varepsilon > 0$ be given, Since $f_n \to f$, $\exists N \in \mathbb{N}$. s.t. $\|f_n - f\|_\infty < \varepsilon$ for all $n \geq N$

Then, for any $n \geq N$, if $|f(\frac{1}{2})| = |\underbrace{f_n(\tfrac{1}{2})}_{0} - f(\frac{1}{2})|$
$\leq \|f_n - f\|_\infty$
$< \varepsilon$

Hence, $f(\frac{1}{2}) = 0$ $\therefore f \in A$ and so A is closed.

(7) Let V be normed Vector spaces, $a \in V$, $r > 0$

$\overline{B_r(a)} := \{x \in V : \|x-a\| \le r\}$ closed ball at a of radius r

Point

Claim : $\overline{B_r(a)}$ is closed why?

Take $(x_n) \subseteq \overline{B_r(a)}$, $x_n \to x \in V$, we show $x \in \overline{B_r(a)}$

$x_n - a \to x - a \overset{HW}{\Longrightarrow} \underbrace{\|x_n-a\|}_{\le r} \to \underbrace{\|x-a\|}_{\le r}$

$\therefore x \in \overline{B_r(a)}$

Pictures

In $\mathbb{R}$: $\overline{B_r(a)} = [a-r, a+r]$

In $\mathbb{R}^2$: $\overline{B_r(a)} =$ 

In $\mathbb{R}^3$: $\overline{B_r(a)} =$ Closed sphere

2.4 Topology 2

Compleeteness and Open sets

Goal 1: Completeness vs Closedness

Prop Let V be normed Vector Spaces, if $C \subseteq V$ is complete, then C is closed

Proof: Suppose C is Complete

Take $(x_n) \subseteq C$ s.t. $x_n \to x \in V$, since (x_n) is Convergent, (x_n) is Cauchy

Thus, $x \in C$, by Completeness ∎ Converse not true

ex) $V = C_{00}$, $C_{00} \subseteq C_{00}$

Goal 2: Introduce open sets

Def'n: Let V be normed Vector Spaces, we say $U \subseteq V$ is open (in V) if $V \setminus U$ is closed

Note: $V \setminus U : \{x \in V : x \notin U\}$

Ex: Let V be normed Vector Spaces, $a \in V$, $r > 0$. $A = \{x \in V : \|x-a\| \geq r\}$ is closed

$\therefore$ $V \setminus A = \{x \in V : \|x-a\| < r\}$ is open

We call $V \setminus A$ the open ball, Centered at a, with radius r, We denote it by $B_r(a)$

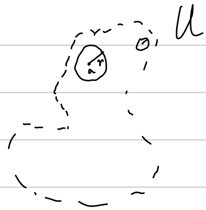
Prop: Let V be normed vector spaces, $U \subseteq V$. the following are equivalent

① U is open in V

② $\forall a \in U, \exists r>0$ s.t. $B_r(a) \subseteq U$

idea

②



Proof: (of Prop)

($\Longrightarrow$): Suppose U is open in V. Thus $V \setminus U$ is closed. Let $a \in U$

For contradiction, suppose for all $r>0$. $B_r(a) \cap (V \setminus U) \neq \emptyset$

For every $n \in \mathbb{N}$, let $x_n \in B_{\frac{1}{n}}(a) \cap (V \setminus U)$

Then, $0 \leq \|x_n - a\| < \frac{1}{n} \longrightarrow 0$ and so $\underset{\in V\setminus U}{x_n} \to \underset{\in U}{a}$

This contradicts that $V \setminus U$ is closed

($\Longleftarrow$) Suppose ②

We will show $V \setminus U$ is closed

Take $(x_n) \subseteq V \setminus U$ s.t. $x_n \to x \in V$

For contradiction, suppose $x \in U$, by ② $\exists r>0$ s.t. $B_r(a) \subseteq U$

For large enough n, $x_n \in B_r(a) \subseteq U$, which is a contradiction

$\therefore x \notin U$ (i.e $x \in V \setminus U$) and so $V \setminus U$ is closed

Def'n: Let V be normed vector spaces, $A \subseteq V$

① We say $x \in V$ is a <u>limit point</u> of A if $\exists (a_n) \subseteq A$ s.t. $a_n \to x$

② We say $x \in A$ is an <u>interior point</u> of A if $\exists r>0$ s.t. $B_r(x) \subseteq A$

Picture: (In $\mathbb{R}^2$)

Summary: Let V be normed vector spaces. $A \subseteq V$

① A is closed in V iff A contains all its limit points

② A is open in V iff every point in A is an interior point of A

3.1 Unions & Intersections

ex): $\underbrace{\bigcap_{i=1}^{\infty} (-\frac{1}{i}, \frac{1}{i})}_{\text{open}} = \underbrace{\{0\}}_{\text{not open}}$

ex) $\underbrace{\bigcup_{i=1}^{\infty} [0, 1-\frac{1}{i}]}_{\text{closed}} = \underbrace{[0, 1)}_{\text{not closed}}$

Prop Let V be normed Vector Spaces

① If $\{A_\alpha\}_{\alpha \in I}$ are open in V, then $\bigcup_{\alpha \in I} A_\alpha$ is open

② If $A_1, A_2, \cdots, A_n$ open in V, then $\bigcap_{i=1}^{n} A_i$ is open

Proof:

① $a \in \bigcup_{\alpha \in I} A_\alpha \Rightarrow \exists \alpha \in I, a \in A_\alpha$

Since A_α is open, $\exists r > 0$ s.t. $B_r(a) \subseteq A_\alpha \subseteq \bigcup_{\alpha \in I} A_\alpha$

② $a \in A_1 \cap A_2 \cap \cdots \cap A_n$, $\forall i$, $a \in A_i$

$\Rightarrow \forall i, \exists r_i > 0, B_{r_i}(a) \subseteq A_i$

Take $r = \min\{r_1, r_2, \cdots, r_n\} \Rightarrow B_r(a) \subseteq A_1 \cap A_2 \cap \cdots \cap A_n$

if has infinite r_i, this will goes wrong, we should write inf

Cor Let V be normed Vector Spaces

① If $\{A_\alpha\}_{\alpha \in I}$ are closed in V, then $\bigcap_{\alpha \in I} A_\alpha$ is closed

② If $A_1, A_2, \cdots, A_n$ are closed in V, then $\bigcup_{i=1}^{n} A_i$ is closed

Why?

$\underbrace{V \setminus (\bigcap_{\alpha \in I} A_\alpha)}_{\text{open}} = \underbrace{\bigcup_{\alpha \in I} (V \setminus A_\alpha)}_{\text{open}}$

$\underbrace{V \setminus (\bigcup_{i=1}^{n} A_i)}_{\text{open}} = \underbrace{\bigcap_{i=1}^{n} (V \setminus A_i)}_{\text{open}}$

3.2 Closures & Interiors

Ideas: Let V be normed Vector Spaces, $A \subseteq V$

① If A is not closed, throw in all the limit points of A to construct a closed set strongly related to A (Closure)

② If A is not open, restrict your attention to only the interior points of A. This gives you an open set strongly related to A (Interior)

Def'n Let V be normed Vector Spaces

① The Closure of A: $\overline{A} = \bigcap_{\substack{A \subseteq C \\ C \text{ closed}}} C$ 包含A的最小闭集

② The interior of A: $\mathrm{Int}(A) = \bigcup_{\substack{\mathcal{U} \subseteq A \\ \mathcal{U} \text{ open}}} \mathcal{U}$ A包含的最大开集

Remark: ① $\overline{A}$ is the smallest closed set in V containing A

② $\mathrm{Int}(A)$ is the largest open set in V contained in A

③ A is closed iff $\overline{A} = A$

④ A is open iff $\mathrm{Int}(A) = A$

$\mathrm{Int}(A) \subseteq A \subseteq B$

$\mathrm{Int}(B) \subseteq B$

Prop: Let V be normed Vector Spaces, $A \subseteq V$

$\overline{A} = \{x \in V : x \text{ limit points of } A\}$

$A \subseteq \overline{A}$ $A \subseteq B$

$B \subseteq \overline{B}$

Proof: Let $X = \{x \in V : x \text{ is limit point of } A\}$

($X \subseteq \overline{A}$) HW

($\overline{A} \subseteq X$)
We must show X is closed and $A \subseteq X$. Indeed, $A \subseteq X$

Claim: X is closed

Let $(x_n) \subseteq X$ s.t. $x_n \to x \in V$

For every $n \in \mathbb{N}$, x_n is a limit point of A and so we may find $y_n \in A$ s.t. $\|y_n - x_n\| < \frac{1}{n}$

Then, $y_n = \underbrace{y_n - x_n}_{\to 0} + \underbrace{x_n}_{\to x}$

$\longrightarrow x$, and so $x \in X$

$\therefore$ X is closed

Prop: Let V be normed vector spaces, $A \subseteq V$

$Int(A) = \{X \in A : X$ is interior point of $A\}$

Proof: Let $X = \{x \in A : x$ is interior point of $A\}$

($Int(A) \subseteq X$) HW

($X \subseteq Int(A)$)

We show X is open and $X \subseteq A$. Obviously, $X \subseteq A$

Claim X is open

Let $x \in X$, thus $\exists\, r > 0$ s.t. $B_r(x) \subseteq A$.

Now, since open ball is open. for all $y \in B_r(x)$, $\exists\, r' > 0$ s.t. $B_{r'}(y) \subseteq B_r(x)$

Thus, $B_{r'}(y) \subseteq A$, and so $y \in X$.

Hence, $B_r(x) \subseteq X$ and so X is open

3.3 Examples

Ex: $A = [0,1)$

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

$\overline{A} = [0,1]$

Ex: The closure of $B_r(a)$ is $\overline{B_r(a)}$

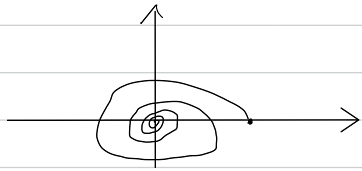
Ex: $\text{Int}(\overline{B_r(a)}) = B_r(a)$

Ex: $A = \mathbb{Q} \subseteq \mathbb{R}$

$\left.\begin{array}{l} \text{Int}(A) = \emptyset \\ \overline{A} = \mathbb{R} \end{array}\right\}$ Math 147

Ex: $A = \{(e^{-x}\cos x, e^{-x}\sin x) : x \geq 0\} \subseteq \mathbb{R}^2$

$\text{Int}(A) = \emptyset$, $\overline{A} = A \cup \{(0,0)\}$



Ex: $V = \ell^\infty$, $C_{00} = \{(x_n) \in V : \text{eventually all } 0\text{'s}\}$, $C_0 = \{(x_n) \in V : x_n \to 0\}$

① show C_0 is closed

Let $(x_n) \subseteq C_0$ s.t. $x_n \to x \in \ell^\infty$

Claim: $x \in C_0$

Say for $n \in \mathbb{N}$, $x_n = (x_n^{(1)}, x_n^{(2)}, \cdots\cdots)$ and $x = (a_1, a_2, \cdots)$

We know for every $n \in \mathbb{N}$, $x_n \in C_0$, and so $x_n^{(k)} \to 0$ as $k \to \infty$

Let $\varepsilon > 0$ be given, we can find $N \in \mathbb{N}$ s.t. $\|x_n - x\|_\infty < \frac{\varepsilon}{2}$, for $n \geq N$

Also, we can find $K \in \mathbb{N}$, s.t. $|x_N^{(k)}| < \frac{\varepsilon}{2}$ for $k \geq K$

Now for $k \geq K$, $|a_k| = |a_k - x_N^{(k)} + x_N^{(k)}| \leq |a_k - x_N^{(k)}| + |x_N^{(k)}| \leq \|x - x_N\|_\infty + |x_N^{(k)}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$

Hence, $x \in C_0$ and so C_0 is closed.

② show $\overline{C_{00}} = C_0$

We have that $C_{00} \subseteq C_0$ and C_0 is closed. Hence, $\overline{C_{00}} \subseteq C_0$

Claim: $C_0 \subseteq \overline{C_{00}}$

Let $x \in C_0$ say $x = (a_1, a_2, \cdots\cdots)$, Hence $a_k \to 0$ as $k \to \infty$

For every $n \in \mathbb{N}$, let $x_n = (a_1, a_2, \cdots, a_n, 0, 0, \cdots) \in C_{00}$

Let $\varepsilon > 0$ be given, we may find $N \in \mathbb{N}$, s.t. $|a_n| < \frac{\varepsilon}{2}$ for $n \geq N$.

For $n \geq N$, $\|x_n - x\|_\infty = \|x - x_n\|_\infty = \|(0, 0, \cdots, a_{n+1}, a_{n+2}, \cdots)\|_\infty \leq \frac{\varepsilon}{2} < \varepsilon$

$\therefore \underset{\in C_{00}}{x_n} \to x \Rightarrow x \in \overline{C_{00}}$ $\quad \therefore C_0 \subseteq \overline{C_{00}} \Rightarrow \overline{C_{00}} = C_0$

3.4 Properties of Closure & Interior

Prop: Let V be normed Vector Spaces, $A, B \subseteq V$

① $Int(A \cup B) \supseteq Int(A) \cup Int(B)$

② $Int(A \cap B) = Int(A) \cap Int(B)$

③ $\overline{A \cup B} = \overline{A} \cup \overline{B}$

④ $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$

Proof: [of ②,③]

② $(\subseteq)$ Let $x \in Int(A \cap B)$

$\Rightarrow \exists r > 0$, $B_r(x) \subseteq A \cap B$

$\Rightarrow B_r(x) \subseteq A$ and $B_r(x) \subseteq B$

$\Rightarrow x \in Int(A)$, $x \in Int(B)$

$\Rightarrow x \in Int(A) \cap Int(B)$

$(\supseteq)$ $\exists r_1, r_2 > 0$ s.t. $B_{r_1}(x) \subseteq A$, $B_{r_2}(x) \subseteq B$

Taking $r = \min\{r_1, r_2\}$, $B_r(x) \subseteq A \cap B \Rightarrow x \in Int(A \cap B)$

③ $(\subseteq)$ Let $x \in \overline{A \cup B}$

$\Rightarrow (x_n) \subseteq A \cup B$, $x_n \to x$

Thus, infinity many $x_n \in A$ or infinity many $x_n \in B$

WLOG, say infinity many $x_n \in A$.

Thus, we may find a subsequence of (x_n), (x_{n_k}) s.t. $(x_{n_k}) \subseteq A$

Also by A1, $\underset{\in A}{x_{n_k}} \to x \Rightarrow x \in \overline{A} \Rightarrow x \in \overline{A} \cup \overline{B}$

$(\supseteq)$ Let $x \in \overline{A} \cup \overline{B}$

WLOG, say $x \in \overline{A} \Rightarrow \exists (x_n) \subseteq A$, $x_n \to x$

$\Rightarrow \exists (x_n) \subseteq A \cup B$, $x_n \to x$

$\Rightarrow x \in \overline{A \cup B}$

ex: $A = [0,1)$, $B = (1,2]$

$\overline{A \cap B} = \overline{\emptyset} = \emptyset$

$\overline{A} \cap \overline{B} = \{1\}$

ex. $A = [0,1]$, $B = [1,2]$

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

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

<u>Prop</u>: Let V be a norm vector space, $A \subseteq V$

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

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

<u>Proof</u>:

we don't need ε! :) Good Proof.

① observe that $A \subseteq \overline{A}$ and so $V \setminus \overline{A} \subseteq V \setminus A$

Since $V \setminus \overline{A}$ is open and $V \setminus \overline{A} \subseteq V \setminus A$, $V \setminus \overline{A} \subseteq \text{Int}(V \setminus A)$

Then, $\text{Int}(V \setminus A) \subseteq V \setminus A$ and so $A \subseteq V \setminus \text{Int}(V \setminus A)$

Since $V \setminus \text{Int}(V \setminus A)$ is closed, $\overline{A} \subseteq V \setminus \text{Int}(V \setminus A)$

$\Longrightarrow \text{Int}(V \setminus A) \subseteq V \setminus \overline{A}$

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

② Letting $B = V \setminus A$,

$\text{Int}(V \setminus B) = V \setminus \overline{B} \Longrightarrow \text{Int}(A) = V \setminus \overline{(V \setminus A)}$

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

4.0 Math 147

Theorm : [Bolzano − Weierstrass]

Every bounded sequence of real numbers has a Convergent subsequence

Def'n : $A \subseteq \mathbb{R}$ is Compact if every sequence $\{a_n\} \subseteq A$ has a subsequence which Converges in A

Remark : If $A \subseteq \mathbb{R}$ is closed and bounded then A is Compact

why? $(a_n) \subseteq A \Rightarrow (a_n)$ is bounded $\xRightarrow{\text{by B-W}} a_{n_k} \to a$

A closed $\Rightarrow a \in A$

Remark: In fact, $A \subseteq \mathbb{R}$ is Compact iff A is closed + bounded. We will Prove the $(\Longrightarrow)$ direction in the next module

4.1 Compactness 1

The idea: Compactness is a topological Property a subset of a NVS can have which makes it "close to" finite

Def'n Let V be a normed vector Space, $A \subseteq V$

We say A is compact if every sequence in A has a subsequence which converges in A

Ex. $A = [0,1) \subseteq \mathbb{R}$

$(1-\frac{1}{n}) \subseteq A$ but $1-\frac{1}{n} \to 1$, every subsequence also converge to $1 \notin A$

$\therefore$ A is not Compact A is not closed

EX. $A = \mathbb{R}$

$(n) \subseteq \mathbb{R}$, every subsequence diverges

$\therefore$ A is not Compact (A is not bounded)

Recall [Math 147]: $A \subseteq \mathbb{R}$ is Compact iff A is closed + bounded

Def'n: Let V be a normed vector Space, $A \subseteq V$

We Say A is bounded if $\exists M > 0$ s.t. $\|a\| \leq M$ for all $a \in A$

Warning! $A \subseteq \mathbb{R}$ Compact iff bounded + closed False! in every NVS

$V = C([0,1])$, $A = \overline{B_1(0)}$. A is closed + bounded

Claim: A is not Compact

$(f_n) \subseteq A$, $f_n(x) = x^n$, by Same argument as A1, every subsequence of $f_n(x)$ diverges

Prop Let V be a normed vector space, $A \subseteq V$
If A is compact then A is closed + bounded

Proof: Suppose A is compact in V

Claim A is closed

Take $(a_n) \subseteq A$ s.t. $a_n \to a \in V$, since A is compact

(a_n) has a subsequence (a_{n_k}) s.t. $a_{n_k} \to b \in A$, by A1, $a = b \in A$

Claim A is bounded

Suppose A is not bounded, for all $n \in \mathbb{N}$ we may find $(a_n) \in A$, s.t. $\|a_n\| \geqslant n$

Then, every subsequence of (a_n) is unbounded and hence divergent.

Contradiction!

4.2 Heine-Borel

Theorm [Heine - Borel]:

A set $A \subseteq \mathbb{R}^n$ is compact iff A is closed + bounded

Recall: [Bolzano - Weierstrass]

Every bounded sequence of real numbers has a convergent subsequence

Conclusion This proves H-B for $\mathbb{R}^1$

Lemma: Let V be a normed vector space, $A \subseteq B \subseteq V$

If A is closed and B is compact then A is compact.

why! $(a_n) \subseteq A \subseteq B$, $a_{n_k} \to b \in B$, A closed $\Rightarrow b \in A$

Lemma $A, B \subseteq \mathbb{R}$, if A, B are compact, then $A \times B \subseteq \mathbb{R}^2$ is compact.

Proof: Suppose $A, B \subseteq \mathbb{R}$ are compact. Let $(a_n, b_n) \subseteq A \times B$ be a sequence

Since $(a_n) \subseteq A$ and A is compact, we may find a subsequence $a_{n_k} \longrightarrow a \in A$

Similarly, $(b_{n_k}) \subseteq B$ must have a subsequence $b_{n_{k_l}} \longrightarrow b \in B$

By A1, $a_{n_{k_l}} \to a$. $\therefore (a_{n_{k_l}}, b_{n_{k_l}}) \to (a, b) \in A \times B$

Hence, $A \times B$ is compact!

Corollary If $A_1, A_2, \cdots, A_n \subseteq \mathbb{R}$ are compact then $A_1 \times A_2 \times \cdots \times A_n \subseteq \mathbb{R}^n$ is compact why? Induction!

Theorm [Heine - Borel]: A set $A \subseteq \mathbb{R}^n$ is compact iff A is closed + bounded

Proof: $\Rightarrow$ Done

$\Leftarrow$ Suppose $A \subseteq \mathbb{R}^n$ is closed and bounded, Since A is bounded, $A \subseteq [-M, M]^n$ for some $M > 0$

By the Corollary, $[-M, M]^n$ is compact. Since A is closed, A is also compact by the lemma.

4.3 Open Covers

Goal: Give an alternate description of Compactness which exhibits the finiteness motivation from Module 1.

Defn Let V be normed vector space, $A \subseteq V$

① An open cover of A is a collection of open set $\{\mathcal{U}_\alpha : \alpha \in I\}$ s.t. $A \subseteq \bigcup_{\alpha \in I} \mathcal{U}_\alpha$

② An open cover $A \subseteq \bigcup_{\alpha \in I} \mathcal{U}_\alpha$ is called finite if $|I| < \infty$

③ A subset of an open cover of A, $\{\mathcal{U}_\alpha : \alpha \in I\}$, which is also an open cover of A, is called a subcover of $\{\mathcal{U}_\alpha : \alpha \in I\}$

Ex. $V = \mathbb{R}$, $A = [0,1]$

An open cover of A: $A \subseteq \bigcup_{\alpha \in [0,1] \cap \mathbb{Q}} (\alpha - \frac{1}{4}, \alpha + \frac{1}{4})$

A finite subcover: $A \subseteq (-\frac{1}{4}, \frac{1}{4}) \cup (0, \frac{1}{2}) \cup (\frac{1}{4}, \frac{3}{4}) \cup (\frac{1}{2}, 1) \cup (\frac{3}{4}, \frac{5}{4})$

Ex. $V = \mathbb{R}^2$, $A = \mathbb{Z} \times \mathbb{Z}$, $A \subseteq \bigcup_{\alpha \in \mathbb{Z} \times \mathbb{Z}} B_{\frac{1}{2}}(\alpha)$

No finite subcover

Ex. $V = \mathbb{R}$, $A = (0,1]$

$A \subseteq \bigcup_{n \in \mathbb{N}} (\frac{1}{n}, 2)$ No finite subcover!

Theorem Let V be normed vector space, $A \subseteq V$

$A \subseteq V$ is compact iff every open cover of A has a finite subcover

4.4 Compactness 2

Lemma Let V be normed vector space, $A \subseteq V$ Compact.

Let $A \subseteq \bigcup_{\alpha \in I} U_\alpha$ be an open Cover of A. There exists $R>0$ s.t. 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 $n \in \mathbb{N}$, $\exists\, a_n \in A$ s.t. $B_{\frac{1}{n}}(a_n) \not\subseteq U_\alpha$ for all $\alpha \in I$

Since $(a_n) \subseteq A$ and A is Compact, there exists $a_{n_k} \to a \in A$

Say $a \in U_{\alpha_0}$, $\alpha_0 \in I$. Pick $M \in \mathbb{N}$ s.t. $B_{\frac{2}{M}}(a) \subseteq U_{\alpha_0}$

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

Then, for $k \geq N$ s.t. $n_k > M$:

Take $x \in B_{\frac{1}{M}}(a_{n_k})$ $\Longrightarrow \|x-a\| = \|x - a_{n_k} + a_{n_k} - a\|$

$$\leq \|x-a_{n_k}\| + \|a_{n_k} - a\|$$

$$< \frac{1}{M} + \frac{1}{M} = \frac{2}{M}$$

$\Longrightarrow x \in B_{\frac{2}{M}}(a)$ $\therefore B_{\frac{1}{M}}(a_{n_k}) \subseteq B_{\frac{2}{M}}(a) \subseteq U_{\alpha_0}$

Since $n_k > M$

Then $B_{\frac{1}{n_k}}(a_{n_k}) \subseteq B_{\frac{1}{M}}(a_{n_k}) \subseteq U_{\alpha_0}$ Contradiction!

Prop [Part 1]. Let V be normed vector space

If $A \subseteq V$ is Compact then every open Cover of A has a finite sub Covers

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 $\exists\, a_1, a_2, \cdots, a_n \in A$ s.t. $A \subseteq B_R(a_1) \cup B_R(a_2) \cup \cdots \cup B_R(a_n)$ we are done

live in U_α

Suppose no such Covering existed: Find $a_1 \in A$

$a_2 \in A$ s.t. $a_2 \notin B_R(a_1)$

$a_3 \in A$ s.t. $a_3 \notin B_R(a_1) \cup B_R(a_2)$

$\vdots$

Since $(a_n) \subseteq A$ and A is Compact, (a_n) has a Converge subsequen

However, for $n<m$, $a_m \notin B_R(a_n) \Rightarrow \|a_n - a_m\| \geq R$ $\therefore$ (a_n) has no Cauchy subsequence $\Rightarrow$ a_n has no Convergent Subsequence.

Contradiction!

5.1 Compactness 3

Prop [Part 2]. Let V be normed vector space

If every open cover of A has a finite subcover, then A is compact

Lemma Let V be normed vector space

Suppose every open cover of A has a finite subcover. If $A \subseteq \bigcup_{\alpha \in I} \mathcal{U}_\alpha$ where each $\mathcal{U}_\alpha$ is relatively open in A

Then $\exists\, \alpha_1, \alpha_2, \cdots, \alpha_n \in I$ s.t. $A \subseteq \mathcal{U}_{\alpha_1} \cup \mathcal{U}_{\alpha_2} \cup \cdots \cup \mathcal{U}_{\alpha_n}$ why?

$A \subseteq \bigcup_{\alpha \in I} \mathcal{U}_\alpha$, $\mathcal{U}_\alpha = A \cap \mathcal{O}_\alpha$, $\mathcal{O}_\alpha \subseteq V$ is open $\Rightarrow A \subseteq \bigcup_\alpha (A \cap \mathcal{O}_\alpha) = A \cap (\bigcup_{\alpha \in I} \mathcal{O}_\alpha) \subseteq \bigcup_\alpha \mathcal{O}$

$\Rightarrow A \subseteq \mathcal{O}_{\alpha_1} \cup \cdots \cup \mathcal{O}_{\alpha_n} \Rightarrow A \subseteq \underbrace{\mathcal{U}_{\alpha_1} \cup \cdots \cup \mathcal{U}_{\alpha_n}}_{\text{finite relatively open}}$

Proof [of Prop]

Suppose $A \subseteq V$ s.t. every open cover of A has a finite subcover.

Consider $(a_n) \subseteq A$, for $k \in \mathbb{N}$, Consider $C_k = \overline{\{a_n : n \geq k\}} \cap A$ (closure), we want to show $\bigcap_{k=1}^{\infty} C_k \neq \emptyset$

each C_k is relatively closed in A. Hence, every $\mathcal{U}_k = A \setminus C_k$ is relatively open in A

For contradiction, assume $\bigcap C_k \neq \emptyset$. then $A = A \setminus \emptyset = A \setminus (\bigcap C_k) = \bigcup (A \setminus C_k) = \bigcup_{k=1}^{\infty} \mathcal{U}_k$. by the lemma $\exists\, i_1 < i_2 < \cdots < i_L$ s.t. $A \subseteq \mathcal{U}_{i_1} \cup \mathcal{U}_{i_2} \cup \cdots \cup \mathcal{U}_{i_L}$

Since $C_1 \supseteq C_2 \supseteq \cdots$, we have $\mathcal{U}_1 \subseteq \mathcal{U}_2 \subseteq \mathcal{U}_3 \subseteq \cdots$

$\therefore A \subseteq \mathcal{U}_{i_L} \subseteq A \Rightarrow A = \mathcal{U}_{i_L}$

$\Rightarrow C_{i_L} = A \setminus \mathcal{U}_{i_L} = A \setminus A = \emptyset$

Hence, $a_{i_L} \in C_{i_L} = \emptyset$ Contradiction!

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

$\therefore$ we may find $n_1 < n_2 < n_3 - \cdots$ s.t. $\|a_{n_k} - a\| < \frac{1}{k}$ $\forall k \in \mathbb{N}$

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

5.2 Limits

<u>Def'n</u> Let V, W be normed vector space, $A \subseteq V$

Let $f: A \to W$ be a function and let $a \in \overline{A \setminus \{a\}}$ limit point of $\overline{A\setminus\{a\}}$

We say the limit of f as x approaches a is $v \in W$

written as $\lim_{x\to a} f(x) = v$

If for all $\varepsilon > 0$, $\exists\, \delta > 0$ s.t. if $x \in A$ with $0 < \|x-a\| < \delta$, then $\|f(x) - v\| < \varepsilon$

<u>Remark</u> why $a \in \overline{A\setminus\{a\}}$??

if $a \notin \overline{A\setminus\{a\}}$ then $\nexists$ (not exist) $x \in A$ s.t. $0 < \|x-a\| < \delta$ for small enough δ

Let V be a normed vector space, $A \subseteq V$, $a \in A$, the following are equivalent

① $a \notin \overline{A\setminus\{a\}}$

② $\exists R > 0$ s.t. $B_R(a) \cap A = \{a\}$ (isolated points)

ex. $A = [0,1] \cup \{2\}$ (isolated point), $f: A \to \mathbb{R}$, $f(x) = x^2$, so $\lim_{x\to 2} f(x)$ does not exists

ex. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2+2y^2} = 0$ [if the domain is unspecificed we assume the domain is wherever the function is defined]

Let $\varepsilon > 0$, choose $\delta = 2\varepsilon$ and suppose $(x,y) \neq (0,0)$ s.t. $0 < \|(x,y) - (0,0)\| < \delta$

$$\left|\frac{xy^2}{x^2+2y^2} - 0\right| = \frac{|x|\cdot y^2}{x^2+2y^2} \le \frac{|x|\cdot y^2}{2y^2} = \frac{|x|}{2} = \frac{\sqrt{x^2}}{2} \le \frac{\sqrt{x^2+y^2}}{2} = \frac{\|(x,y)\|}{2} < \frac{2\varepsilon}{2} = \varepsilon$$

ex) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{2xy}{x^2+y^2}$ does not exist

Assume $\displaystyle\lim_{(x,y)\to(0,0)} \frac{2xy}{x^2+y^2} = v \in \mathbb{R}$

As $t \to 0$, $(0,t) \to (0,0)$ and so $\displaystyle\lim_{t\to 0} \frac{2\cdot 0\cdot t}{0^2+t^2} = 0 = v$

as $t \to 0$, $(t,t) \to (0,0)$ $\therefore \displaystyle\lim_{t\to 0} \frac{2t^2}{t^2+t^2} = 1 = v$ Contradiction!

Hence, the limit does not exises

5.3 Continuity

<u>Def'n</u> V, W be normed vector space, $A \subseteq V$

we say $f: A \to W$ is continuous at $a \in A$

If $\forall \varepsilon > 0$ $\exists \delta > 0$ s.t. if $x \in A$ with $0 < \|x-a\| < \delta$ then $\|f(x) - f(a)\| < \varepsilon$

we say f is continuous if f is continuous at every $a \in A$

Remark : $f: A \to W$ $A \subseteq V$ suppose $a \in A$ with $a \notin \overline{A \setminus \{a\}}$, then $\exists r > 0$

s.t. $B_r(a) \cap A = \{a\}$ (i.e. a is an isolated points of A)

Let $\varepsilon > 0$, then ....

choose $\delta = r$, then if $x \in A$ with $\|x - a\| < \delta$, then $x = a$

Thus, $\|f(x) - f(a)\| = \|f(a) - f(a)\| = 0 < \varepsilon$

$\therefore$ f is continuous at a always has continuity at isolated point a

Remark: $f: A \to W$ $A \subseteq V$

Now assume $a \in \overline{A \setminus \{a\}}$, then by defn f is continuous at $a \in A$ iff $\lim_{x \to a} f(x) = f(a)$

Summary $f: A \to W$ continuous iff $\lim_{x \to a} f(x) = f(a)$ for all $a \in A$ with $a \in \overline{A \setminus \{a\}}$

Prop : $f: A \to W$, $A \subseteq V$ and $a \in A$, the following are equivalent

① f is continuous at $a \in A$

☆ ② If $(a_n) \subseteq A$ with $a_n \to a$, then $f(a_n) \to f(a)$

③ If $U \subseteq W$ is open then $f^{-1}(U)$ is relatively open in A

<u>Proof</u> : Prop ①⟺② and A2 ②⟺③

<u>Proof</u> : ①→②. suppose f is continuous at $a \in A$

Let $(a_n) \subseteq A$ s.t. $a_n \to a$

<u>Claim:</u> $f(a_n) \to f(a)$, Let $\varepsilon > 0$ be given, we know $\delta > 0$ if $x \in A$ with $\|x - a\| < \delta$ then $\|f(x) - f(a)\| < \varepsilon$

Chose $N \in \mathbb{N}$ s.t. $\|a_n - a\| < \delta$ for all $n \geq N$. For $n \geq N$, $\|a_n - a\| < \delta \Rightarrow \|f(a_n) - f(a)\| < \varepsilon$

②→① Suppose ②

Assume is not continuous at a $\therefore \exists \varepsilon > 0$ and $a_n \in A$ $(n \in \mathbb{N})$ s.t. $\|a_n - a\| < \frac{1}{n}$ but $\|f(a_n) - f(a)\| \geq \varepsilon$

Thus, $a_n \to a$ and so $f(a_n) \to f(a)$ Contradiction!

Prop $f, g: A \to W$ Continuous, $A \subseteq V$

① $f+g$ is Continuous

② $\forall a \in \mathbb{R}$, $a \cdot f$ is Continuous

③ $f \cdot g$ is Continuous

④ $\frac{f}{g}$ is Continuous provided $g(x) \neq 0$ for all $x \in A$

Why?

$(a_n) \subseteq A$, $a_n \to a \in A$ we have $f(a_n) \to f(a)$ and $g(a_n) \to g(a)$

by limit laws: $f(a_n) + g(a_n) \to f(a) + g(a)$

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

Prop If f, g are Continuous with $f \circ g$ defined, then $f \circ g$ is Continuous

Why? $a_n \to a$, $g(a_n) \to g(a)$ (Continuity of g)

$f(g(a_n)) \to f(g(a))$ (Continuity of f)

5.4 Compactness & Continuity

By A2. Prop V, W are normed vector space, $C \subseteq V$ compact (not empty)

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

Theorem (EVT): V is normed vector space, $C \subseteq V$ compact (not empty)

If $f: C \to \mathbb{R}$ is continuous, then $\exists\ a, b \in C$ s.t $f(a) \leq f(x) \leq f(b)$ for all $x \in C$

Proof: by Prop $f(C)$ is compact, so $f(C)$ is closed and bounded

Consider $y_1 = \inf(f(C)) < \infty$ $\quad y_2 = \sup(f(C)) < \infty$

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

$\therefore\ \exists\ a, b \in C$ s.t. $y_1 = f(a)$, $y_2 = f(b)$

$\Rightarrow f(a) = \min f(C)$, $f(b) = \max f(C)$

s.t. $f(a) \leq f(x) \leq f(b)$ for all $x \in C$

Remark V, W are normed vector space, $K \subseteq V$ compact

$C(K, W) = \{f: K \to W \mid f \text{ continuous}\}$ is a NVS when equipped with the uniform norm

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

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

Why? $f: K \to W$ continuous $\quad \|\cdot\|: W \to \mathbb{R}$ continuous

$\Rightarrow \|\cdot\| \circ f$ continuous

$(\|\cdot\| \circ f)(x) = \|f(x)\|$

Remark If $W = \mathbb{R}$ we write $C(K)$ instead of $C(K, W)$

6.1 Uniform Continuity

Let V, W be normed vector spaces, $A \subseteq V$, $f: A \to W$

Recall: f is continuous iff $\forall a \in A$, $\forall \varepsilon > 0$, $\exists \delta > 0$, $\forall x \in A$ $\|x-a\| < \delta \Rightarrow \|f(x) - f(a)\| < \varepsilon$

Def'n: f is uniformly continuous iff $\forall \varepsilon > 0$, $\forall a, b \in A$, $\|a-b\| < \delta \Rightarrow \|f(a) - f(b)\| < \varepsilon$

Big idea: The δ works uniformly for continuous at a, for all $a \in A$
same delta

Remark: Uniformly continuous $\Rightarrow$ continuous

ex) $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ show f is not uniformly continuous

Suppose f is uniformly continuous, for $\varepsilon = 1$, $\exists \delta > 0$ s.t. $a, b \in \mathbb{R}$ with $|a-b| < \delta$ then $|a^2 - b^2| < 1$

Take $N \in \mathbb{N}$ s.t. $\frac{1}{N} < \delta$

In particular, $|N + \frac{1}{N} - N| < \delta$, and so $|(N+\frac{1}{N})^2 - N^2| < 1$

$$\Rightarrow N^2 + 2 + \frac{1}{N^2} - N^2 < 1$$

$$\Rightarrow 2 + \frac{1}{N^2} < 1$$

$$\Rightarrow 2 < 1$$ Contradiction

ex) $f: (0, 2) \to \mathbb{R}$, $f(x) = \ln(x)$, show f is not uniformly continuous

Suppose f is uniformly continuous, let $\varepsilon = 1$ so that $\exists \delta > 0$ s.t. if $(a, b) \in (0, 2)$ with $|a-b| < \delta$ then $|\ln(a) - \ln(b)| < 1$

Take $N \in \mathbb{N}$ s.t. $\frac{1}{N} < \delta$, we have $\frac{1}{N} - \frac{1}{N^2} < \frac{1}{N} < \delta$ $\Rightarrow \ln(\frac{1}{N}) - \ln(\frac{1}{N^2}) < 1$

$$\Rightarrow \ln(N) < 1$$

for large enough N, this is a contradiction

6.2 Compactness Proof

<u>Theorem</u> : $C \subseteq V$ is Compact, if $f: C \to W$ Continuous, then f is Uniformly Continuous

<u>Proof</u> : Suppose $f: C \to W$ is Continuous but not Uniformly Continuous, therefore $\exists\, \varepsilon > 0$ and $(a_k), (b_k) \subseteq C$ s.t. $\|a_k - b_k\| < \frac{1}{k}$, $\|f(a_k) - f(b_k)\| \geq \varepsilon$

Since C is Compact, $\exists\, a_{k_\ell} \to a \in C$. Note $b_{k_\ell} = \underbrace{a_{k_\ell}}_{\to a} + \underbrace{b_{k_\ell} - a_{k_\ell}}_{\to 0} \to a + 0 = a$

Since f is Continuous, $f(a_{k_\ell}) \to f(a)$, $f(b_{k_\ell}) \to f(a)$ $\implies \|f(a_{k_\ell}) - f(b_{k_\ell})\| \to 0$ Contradiction

6.3 Uniform Convergence

Space of functions and uniform Convergence

Idea: We now focus on NVS's which Consist of functions eg) $C(K,W)$

Notation: V, W, NVS, $A \subseteq V$

Q: Let (f_n) be a sequence of function from $A \to W$ what should it mean for (f_n) to "Converge" to some $f: A \to W$?

Two ideas: ① Pointwise Convergence

② Uniform Convergence

Remark: Here, we are not claiming the f_n's and f belong to a particular NVS

Notation: $f, g: A \to W$, $\|f-g\|_\infty = \sup\{\|f(x)-g(x)\| : x \in A\}$

Remark: we may have $\|f\|_\infty = \infty$

eg) $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x$

Idea We are borrowing norm-like notation to talk about the "distance" between f and g

Recall If $f \in C(K,W)$, where $K \subseteq V$ is Compact, then $\|f\|_\infty < \infty$. In fact we know $(C(K,W), \|\cdot\|_\infty)$ is a NVS

Defn $(f_n): A \to W$, $f: A \to W$

① We say (f_n) Converges to f Pointwise (written $f_n \to f$ Pointwise) if $\forall x \in A$, $f_n(x) \to f(x)$

② We say (f_n) Converges to f uniformly (written $f_n \to f$ uniformly) if $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$, $\forall x \in A$: $n \geq N \Rightarrow \|f_n(x) - f(x)\| < \varepsilon$

Idea: The same N works uniformly for all $x \in A$ to tell us that $\|f_n(x) - f(x)\| < \varepsilon$ for $n \geq N$

Remark: $\forall x \in A$ $\|f(x)-g(x)\| < \varepsilon \iff \|f-g\|_\infty < \varepsilon$ (important)

Therefore, $f_n \to f$ uniformly iff ① $\|f_n - f\| < \infty$ eventually $(n \geq M)$

② $\|f_n - f\| \to 0$ as $n \to \infty$

6.4 Examples

ex) $f_n : [0,1] \to \mathbb{R}$, $f_n(x) = x^n$

For $x \in [0,1]$, $f_n(x) \to f(x)$ where $f(x) = \begin{cases} 0 & , x \in [0,1) \\ 1 , & x = 1 \end{cases}$

$\therefore$ $f_n \to f$ Pointwise

Now, $\|f_n - f\|_\infty = 1 \not\to 0$, thus the convergence is not uniform

ex) $f_n : (0,\infty) \to \mathbb{R}$, $f_n(x) = \frac{nx}{e^{nx}}$

$\lim_{n\to\infty} \frac{nx}{e^{nx}} \overset{L'H}{=} \lim_{n\to\infty} \frac{x}{x \cdot e^{nx}} = \lim_{n\to\infty} \frac{1}{e^{nx}} = 0$ $\therefore$ $f_n \to 0$ Pointwise

Now, $\|f_n - 0\|_\infty = \sup\{ |\frac{nx}{e^{nx}}| : x > 0 \} = \sup\{ \frac{nx}{e^{nx}} : x > 0 \} \geq \frac{n \cdot \frac{1}{n}}{e^{n \cdot \frac{1}{n}}} = \frac{1}{e}$

$\therefore$ $\|f_n - 0\|_\infty \not\to 0$ $\Rightarrow$ Convergence is not uniform

ex) $f : [0,1] \times [0,1] \to \mathbb{R}^2$

$f_n(x,y) = (\frac{x+y}{n} , \frac{\sin(xy)}{x^2+n^2})$

for $(x,y) \in [0,1] \times [0,1]$, $\lim_{n\to\infty} f_n(x,y) = 0$ $\therefore$ $f_n \to 0$ Pointwise

Now, $\|f_n - 0\|_\infty = \sup\{ \|(\frac{x+y}{n} , \frac{\sin xy}{x^2+n^2})\| : (x,y) \in [0,1] \times [0,1] \}$

$= \sup\{ \sqrt{(\frac{x+y}{n})^2 + (\frac{\sin(xy)}{x^2+n^2})^2} : (x,y) \in [0,1] \times [0,1] \}$

$\leq \sqrt{(\frac{2}{n})^2 + (\frac{1}{n^2})^2}$

$= \sqrt{\frac{4}{n^2} + \frac{1}{n^4}} \to 0$ as $n \to \infty$

$\therefore$ $f_n \to f$ uniformly

6.5 Two Theorems

Theorem: (f_n) $A \to W$ Continuous if $f_n \to f$ uniformly then f is continuous

Proof Suppose each f_n is continuous and $f_n \to f$ uniformly

Take $(a_n) \subseteq A$ s.t. $a_n \to a \in A$

Claim: $f(a_n) \to f(a)$

Let $\varepsilon > 0$ be given

Since $f_n \to f$ uniformly, $\exists N_1 \in \mathbb{N}$ s.t. $\|f_k - f\|_\infty < \frac{\varepsilon}{3}$ for $k \geq N_1$

Moreover, Since f_{N_1} is Continuous, $\exists N_2 \in \mathbb{N}$ s.t. $\|f_{N_1}(a_n) - f_{N_1}(a)\| < \frac{\varepsilon}{3}$ for all $n \geq N_2$

Then, for $n \geq N_2$, $\|f(a_n) - f(a)\| \leq \|f(a_n) - f_{N_1}(a_n)\| + \|f_{N_1}(a_n) - f_{N_1}(a)\| + \|f_{N_1}(a) - f(a)\|$

$$\leq \|f - f_{N_1}\|_\infty + \|f_{N_1}(a_n) - f_{N_1}(a)\| + \|f_{N_1} - f\|_\infty$$

$$< \varepsilon$$

$\therefore$ $f(a_n) \to f(a)$ $\Longrightarrow$ f is continuous as required

ex) $f_n : [0,1] \to \mathbb{R}$, $f_n(x) = x^n$, $f(x) = \begin{cases} 0, & x \in [0,1) \\ 1, & x = 1 \end{cases}$, $f_n \to f$ Pointwise not hold Continuous

Theorem $K \subseteq V$ Compact, W Banach Space, then $C(K,W)$ is a Banach space

Proof Let $(f_n) \subseteq C(K,W)$ be Cauchy, take $x \in K$ Consider $(f_n(x)) \subseteq W$

Claim: $(f_n(x))$ is Cauchy

Let $\varepsilon > 0$, $\exists N \in \mathbb{N}$ s.t. $\|f_n - f_m\|_\infty < \varepsilon$ for $n, m \geq N$

Then for $n, m \geq N$, $\|f_n(x) - f_m(x)\| \leq \|f_n - f_m\|_\infty < \varepsilon$, this Prove the claim

Since W is Banach space, $f_n(x) \to f(x) \in W$, by doing this for all $x \in K$,

We have created a function $f: K \to W$ s.t. $f_n \to f$ Pointwise

Claim: $f_n \to f$ uniformly

Let $\varepsilon > 0$, $\exists M \in \mathbb{N}$ s.t. if $n, m \geq M$, then $\|f_n - f_m\|_\infty < \frac{\varepsilon}{2}$

Let $n \geq M$ and $x \in K$, then $\|f_n(x) - f(x)\| = \lim_{m \to \infty} \underbrace{\|f_n(x) - f_m(x)\|}_{< \frac{\varepsilon}{2}}$

$$\leq \frac{\varepsilon}{2} < \varepsilon$$

$\therefore$ $\|f_n - f\|_\infty \leq \frac{\varepsilon}{2} < \varepsilon$ for all $n \geq M$, this Proved the claim

By Previous Theorem, $f \in C(K,W)$, Hence $f_n \to f$ in $C(K,W)$ $\blacksquare$

Week 7

Welcome to Part 2 of the course! This week we will be starting our study of multivariable calculus. Naturally, we will start by investigating multivariable differentiation. As you will soon see, we will consistently use the language of real analysis in our theory. So, don't go forgetting all of the real analysis you just learned!

1 Partial Derivatives

In this section we will be exploring partial derivatives and the differentiability of multivariable functions. We will freely use the theory of differentiability for functions $f : A \to \mathbb{R}$, $A \subseteq \mathbb{R}$.

Definition. Let $f : A \to \mathbb{R}$, $A \subseteq \mathbb{R}^n$, be a multivariable function. We denote such a function by $f(x_1, x_2, \ldots, x_n)$. Fix $1 \leq i \leq n$. We define the **partial derivative** of f, with respect to x_i, at $a = (a_1, a_2, \ldots, a_n) \in A$ to be

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

provided the limit exists. Here e_i is the i^{th} standard basis vector of $\mathbb{R}^n$.

Notation. We also denote $\frac{\partial f}{\partial x_i}(a)$ by $f_{x_i}(a)$.

The big picture here is that the partial derivative of f, w.r.t x_i, at a is obtained by differentiating the function with respect to the variable x_i, while treating all x_j $(i \neq j)$ like constants, and then plugging in the point a. As usual, we denote by $\frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i} f(x_1, x_2, \ldots, x_n)$ the function which takes in a point in A and then gives the partial derivative of f w.r.t. x_i at that point, provided it exists.

Example. *Consider $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = e^x \cos(y) + x^2 y$. Then,*

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

and

$$\frac{\partial}{\partial y} f(x, y) = -e^x \sin(y) + x^2.$$

We now use these partial derivatives to define the more general multivariable partial derivative.

Definition. Let $f : A \to \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$, be a function. There exist real valued functions functions $f_1, f_2, \ldots, f_m$ such that $f = (f_1, f_2, \ldots, f_m)$. For $1 \leq i \leq n$, we define the **partial derivative** of f, w.r.t. x_i, at $a \in A$ by

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

provided each $\frac{\partial f_j}{\partial x_i}$ exists.

Example. *Consider $f : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $f(x, y, z) = (xy, e^{xz})$. Here $f_1(x, y, z) = xy$ and $f_2(x, y, z) = e^{xz}$. Moreover, for any $(x, y, z) \in \mathbb{R}^3$,*

$$f_x(x, y, z) = (y, ze^{xz}),$$

$$f_y(x, y, z) = (x, 0),$$

and

$$f_z(x, y, z) = (0, xe^{xz}).$$

2 Differentiability

Goal: We will now define the notion of differentiability at a point for functions $f : A \to \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$. We will then discuss how this definition of differentiability relates to the partial derivatives already established.

Recall. *(Linear Approximation Theorem)* Remember that a function $f : A \to \mathbb{R}$, $A \subseteq \mathbb{R}$, is differentiable at a point $a \in A$ if and only if f is defined on an open interval containing a and there exists a function $T(x) = mx$ such that

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

This above recall/theorem will be our motivating piece for what follows.

Notation. We let $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ denote the set of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Definition. Let $a \in A \subseteq \mathbb{R}^n$ and $f : A \to \mathbb{R}^m$ be a function. We say f is **differentiable** at a if there exists an open set $a \in U \subseteq A$ such that f is defined on U and there exists $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ such that

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

$$\frac{f(a+h) - f(a) - T(h)}{h} = 0$$

Remark.

1. In the above definition, we assume h is small enough that $a + h \in U$, so that $f(a+h)$ is for sure defined.

2. Notice that $T \in \mathcal{L}(\mathbb{R}, \mathbb{R})$ if and only if $T(x) = mx$, so this definition coincides with our previous characterization.

To prove our next theorem, we need some help from your third assignment!

Definition. Let A be an $m \times n$ matrix of real numbers. We define

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

to be the **operator norm** on the set of all $m \times n$ real matrices, $M_{m\times n}(\mathbb{R})$.

One quick thing to note is that if $A \in M_{m\times n}(\mathbb{R})$ and $0 \neq x \in \mathbb{R}^n$, then the norm of $x/||x||$ is 1 and so $||A(x/||x||)|| \leq ||A||_{op}$. Hence, $||Ax|| \leq ||A||_{op} \cdot ||x||$. Notice that this final inequality trivially holds when $x = 0$.

Theorem. Let $a \in A \subseteq \mathbb{R}^n$ and $f : A \to \mathbb{R}^m$ be a function. If f is differentiable at a then f is continuous at a.

Proof. Suppose f is differentiable at $a \in A$ so that f is defined on an open set $a \in U \subseteq A$ and there exists $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ such that

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

Now, let $B \in M_{m \times n}(\mathbb{R})$ be the standard matrix of T (ie. relative to the standard bases for $\mathbb{R}^n$ and $\mathbb{R}^m$). Recall from linear algebra that this means $T(x) = Bx$ for all $x \in \mathbb{R}^n$. In particular, there exists $\delta > 0$ such that for $0 < ||h|| < \delta$,

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

Therefore, for $0 < ||h|| < \delta$,

$$\begin{aligned} &||f(a+h) - f(a) - Bh|| < ||h||, \\ &\implies ||f(a+h) - f(a)|| - ||Bh|| < ||h|| \text{ (Reverse Triangle Ineq.)} \\ &\implies ||f(a+h) - f(a)|| < ||Bh|| + ||h|| \leq ||B||_{op}||h|| + ||h||. \end{aligned}$$

As $h \to 0$ in $\mathbb{R}^n$, $||B||_{op}||h|| + ||h|| \to 0$ in $\mathbb{R}$, and so $||f(a+h) - f(a)|| \to 0$ by the Squeeze Theorem. In particular,

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

Letting $x = a + h$, we see that

$$\lim_{x \to a} f(x) = f(a).$$

□

3 Partial Derivatives AND Differentiability

Goal: Determine a relationship between differentiability and partial derivatives.

Theorem. Let $a \in A \subseteq \mathbb{R}^n$ and $f : A \to \mathbb{R}^m$ be a function. If f is differentiable at a then all the partial derivatives of f exist at a.

Proof. Suppose f is differentiable at a so that there exists an open set $a \in U \subseteq A$ such that f is defined on U and there exists $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ such that

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

Letting B be the standard matrix of T, we see that

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

Let $\{e_1, e_2, \ldots, e_n\}$ be the standard basis for $\mathbb{R}^n$. Consider $h = te_i$, where $t \in \mathbb{R}$. Observe that $t \to 0$ if and only if $h \to 0$. Moreover,

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

and so

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

Similarly,

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

and

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

Therefore,

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

and so,

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

□

Remark. *(Super Important!)* With notation as above, we have just proved that

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

By matrix multiplication, Be_i is actually just the i^{th} column of B. In particular, if $B = [b_{i,j}]$ then $Be_i = (b_{1,i}, b_{2,i}, \ldots, b_{m,i})$. Moreover, letting $f = (f_1, f_2, \ldots, f_m)$ as usual, we have that

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

and so

$$b_{i,j} = \frac{\partial f_i}{\partial x_j}(a).$$

This tells us what the matrix B actually is, opposed to just a theoretical matrix that exists. The below definition gives this matrix a special name.

Definition. If all first order partial derivatives of $f : A \to \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$, exist at $a \in A$, 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. The above Theorem tells us that if f is differentiable at a, then the total derivative $B = Df(a)$ exists and strongly relates to the definition of differentiability.

Definition. When $m = 1$, $Df(a) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right)$ is called the **gradient** of f at a and is labelled by $\nabla f(a)$.

We conclude this lecture with some examples.

Example. *Consider $A = \{(x, y) \in \mathbb{R}^2 : x < y\}$ and $f : A \to \mathbb{R}^3$ given by*

$$f(x, y) = \left(\sqrt{y - x}, xy + 2, \ln(y^3 - x^3 + 1)\right).$$

Then,

$$Df(x, y) = \begin{bmatrix} \frac{-1}{2\sqrt{y-x}} & \frac{1}{2\sqrt{y-x}} \\ y & x \\ \frac{-3x^2}{y^3-x^3+1} & \frac{3y^2}{y^3-x^3+1} \end{bmatrix}.$$

Notice that here I am giving the total derivative as a function of the points in A. I can do this because the partial derivatives of f all exist at ***every*** *point in A.*

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

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

We first find $\nabla f(0,0)$. Note that

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f((0,0)+he_1) - f(0,0)}{h} = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = 0,$$

and $\frac{\partial f}{\partial y}(0,0) = 0$, similarly. Hence $\nabla f(0,0) = (0,0)$. ***In particular, the total derivative of f exists at $(0,0)$.*** *However, $(1/n, 1/n) \to (0,0)$ but $f(1/n,1/n) \to 1/2 \neq f(0,0)$.* ***Therefore f is not continuous, and hence not differentiable, at $(0,0)$.*** *So what have we learned? Well, a total derivative existing at a point does NOT mean f is differentiable at that point! That is, don't use the (false) converse of the above Theorem!*

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

Proof. Suppose every partial derivative of f exists on A and that every partial derivative is continuous at a.
Case 1: $m = 1$.
Suppose $a = (a_1, a_2, \ldots, a_n)$. Since A is open there exists $r > 0$ such that $B_r(a) \subseteq A$. 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 $1 \le j \le n$ 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 on A, as $h \to 0$, each $\delta_j \to 0$, and so $\delta \to 0$ in $\mathbb{R}^n$. Therefore

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

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.

Case 2: Follows from Case 1 and the fact that $f = (f_1, f_2, \ldots, f_m)$. □

4 Examples

1. Let $A = \{(x, y) : x > 0, y > 0\}$, which is open in $\mathbb{R}^2$. Consider $f : A \to \mathbb{R}^2$ given by $f(x, y) = \left(\sqrt{x + 2y}, \frac{\sin(xy)}{x}\right)$. Prove that f is differentiable on A (ie. at every point in A).

   Proof. We see that

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

   and

   $$f_y(x, y) = \left(\frac{1}{\sqrt{x + 2y}}, \frac{x\cos(xy)}{x}\right)$$

   exist on A. Moreover, both f_x and f_y are continuous on A so that by our theorem, f is differentiable at every point in A.

   $\square$

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

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

   Is f differentiable at $(0, 0)$?

   Proof. We see 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 $(0, 0)$. However, this doesn't mean that f is not differentiable at $(0, 0)$. It just means that we can't use our second theorem. We are then left to go back to the definition of differentiability and see if

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

Well,

$$f_x(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} h^2 \sin\left(\frac{1}{|h|}\right) = 0,$$

by a standard squeeze theorem argument. Similarly, $f_y(0,0) = 0$, and so $Df(0,0) = (0,0)$. Now,

$$\begin{aligned}
&\lim_{h\to 0} \frac{f(0+h) - f(0) - Df(0,0)h}{||h||} = \lim_{h\to 0} \frac{f(h)}{||h||} \\
&= \lim_{(h_1,h_2)\to(0,0)} \sqrt{h_1^2 + h_2^2} \sin\left(\frac{1}{\sqrt{h_1^2 + h_2^2}}\right) \\
&= \lim_{(x,y)\to(0,0)} \sqrt{x^2 + y^2} \sin\left(\frac{1}{\sqrt{x^2 + y^2}}\right) \\
&= 0,
\end{aligned}$$

again by the squeeze theorem (real-valued limits preserve order). Therefore f is indeed differentiable at $(0,0)$, even though things were looking bad! $\square$

3. Consider $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x,y) = \dfrac{x^2y + x}{x^2 + y^2}$ if $(x,y) \neq (0,0)$ and $f(0,0) = 0$. Is f differentiable at $(0,0)$?

Proof. Note that

$$f_x(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{h}{h^2},$$

which doesn't exist. By the first theorem of section 3, f is not differentiable at $(0,0)$. $\square$

8.1 Tangent Hyperplanes

Goal: understand the geometrical interpretation of the total derivative (gradient) of scalar functions $f: U \to \mathbb{R}$ where $U \in \mathbb{R}^n$ is open

Motivation

$n=1$ $U \subseteq \mathbb{R}$ open, if $f: U \to \mathbb{R}$ is differentiable at $a \in U$, then $f'(a) = \nabla f(a)$ is the slope of the tangent line to the curve $y = f(x)$ at $x = a$

$n=2$ $U \subseteq \mathbb{R}^2$ open

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

Def$_1$ A hyperplane in $\mathbb{R}^n$ is a set of the form $P = \{(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n : a_1x_1 + a_2x_2 + \cdots + a_nx_n = d\}$ for some fixed $a_1, a_2, \cdots, a_n \in \mathbb{R}$ (not all zero) and $d \in \mathbb{R}$

Remark

$n=2$ hyperplanes = lines

$n=3$ hyperplanes = planes

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

Def$_2$ Let $P = \{(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n : a_1x_1 + a_2x_2 + \cdots + a_nx_n = d\}$ be a hyperplane in $\mathbb{R}^n$, we call $n = (a_1, a_2, \cdots, a_n)$ the normal (vector) of P

Geometrically

Let $b = (b_1, b_2, \cdots, b_n) \in \mathbb{R}$, then $d = a_1b_1 + a_2b_2 + \cdots + a_nb_n$

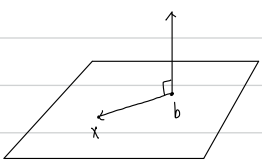
$\therefore$ $x = (x_1, x_2, \cdots, x_n) \in P$

$\iff d = a_1x_1 + a_2x_2 + \cdots + a_nx_n$

$\iff 0 = d - d$

$= a_1(x_1 - b_1) + \cdots + a_n(x_n - b_n)$

$= n \cdot (x - b)$ dot product



$P \subseteq \mathbb{R}^3$

$\therefore$ $P = \{x \in \mathbb{R}^n : n \cdot (x - b) = 0\}$

i.e. $x \in P$ iff n is orthogonal / perpendicular to $x - b$

Def'n $A \subseteq \mathbb{R}^n$, $a \in A$, A hyper plane $a \in P \subseteq \mathbb{R}^n$ with normal n is said to be tangent to A at a if

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

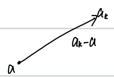
for all sequence $(a_k) \subseteq A \setminus \{a\}$ s.t. $a_k \to a$

Why is this a good def'n ?

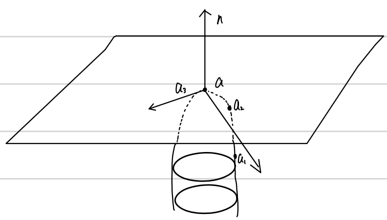
Recall that $a, b \in \mathbb{R}^n$ are orthogonal (perpendicular) iff $a \cdot b = 0$

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

says that unit (length 1) vectors in the direction



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



Theorem $\mathcal{U} \subseteq \mathbb{R}^n$ open, $a \in \mathcal{U}$ $f: \mathcal{U} \to \mathbb{R}$, if f is differentiable at a then the surface $S = \{(x, z) \in \mathbb{R}^{n+1} : z = f(x), x \in \mathcal{U}\}$ has a tangent hyperplane at $(a, f(a))$ with normal $n = (\nabla f(a), -1)$

Proof See the Appendix

ex) Find the tangent plane to the surface at $(1, 1, 3)$

Consider $f: \mathbb{R}^2 \to \mathbb{R}$ with $f(x,y) = 2x^2 + y^2$ Note that f_x, f_y exist are continuous on $\mathbb{R}^2$

$\therefore f$ is differentiable on $\mathbb{R}^2$

$\nabla f(x,y) = (4x, 2y)$, $\nabla f(1,1) = (4, 2)$, $n = (4, 2, -1)$

$4x + 2y - z = d \Rightarrow d = 4 \cdot 1 + 2 \cdot 1 - 3 = 3$

Then, the tangent plane is $4x + 2y - z = 3$ at $(1, 1, 3)$

8.2 Basic Properties

Theorem $A \subseteq \mathbb{R}^n$, $f, g : A \to \mathbb{R}^m$, $a \in A$, if f and g are diff at a

Then ① $f+g$ is diff at a and $D(f+g)(a) = Df(a) + Dg(a)$

② $\forall\, \alpha \in \mathbb{R}$, αf is diff at a and $D(\alpha f)(a) = \alpha Df(a)$

③ The function $f \cdot g : A \to \mathbb{R}$, $(f \cdot g)(x) = f(x) \cdot g(x)$ is diff at a

dot Product

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

matrix multiplication + addition

Proof: ①, ② Piazza discussion

③ Appendix

Theorem (Chain rule)

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

If f is diff at $a \in A$ and g is diff at $f(a) \in B$, then $g \circ f$ is diff at a with $D(g \circ f)(a) = Dg(f(a))\, Df(a)$

Proof: Appendix

General Proof Strategy

To show $Df(a) = X$, we must show $\lim\limits_{h \to 0} \dfrac{f(a+h) - f(a) - Xh}{\|h\|} = 0$

Remark: $f(u_1, u_2, \cdots, u_m)$ diff $\mathbb{R}^m \to \mathbb{R}$

$g(x_1, x_2, \cdots, x_n)$ diff $\mathbb{R}^n \to \mathbb{R}^m$

$g = (g_1, g_2, \cdots, g_m)$

By the chain Rule $\nabla(f \circ g)(x_1, x_2, \cdots, x_n) = \nabla f(g(x_1, x_2, \cdots, x_n)) \cdot Dg(x_1, x_2, \cdots, x_n)$

For $1 \le k \le m$ let $u_k = g_k(x_1, x_2, \cdots, x_n)$

By Comparing the i^{th} entries of ★: if $z = f(g(x_1, x_2, \cdots, x_n))$ $\dfrac{\partial z}{\partial x_i} = \dfrac{\partial f}{\partial u_1}\dfrac{\partial u_1}{\partial x_i} + \dfrac{\partial f}{\partial u_2}\dfrac{\partial u_2}{\partial x_i} + \cdots + \dfrac{\partial f}{\partial u_m}\dfrac{\partial u_m}{\partial x_i}$

8.3 Mean Value Theorem

Recall MVT in $\mathbb{R}$

If $f:[a,b]\to\mathbb{R}$ is Continuous and diff on (a,b) then $\exists\, c\in(a,b)$ s.t. $f(b)-f(a)=f'(c)(b-a)$

Guess If $f:\mathbb{R}^n\to\mathbb{R}^m$ is diff and $a,b\in\mathbb{R}^n$ then $\exists\, c$ "between" a,b s.t. $f(a)-f(b)=Df(c)(b-a)$

Problem 1 What should between $a,b\in\mathbb{R}^n$ mean

Answer On the Line Segment

$L(a,b):=\{(1-t)a+tb : t\in[0,1]\}$

Problem 2

$f\ \mathbb{R}\to\mathbb{R}^2$, $f(x)=(\cos x,\sin x)$

$f(0)=f(2\pi)$ Does there exist $c\in[0,2\pi]$ s.t. $0=Df(c)\cdot 2\pi$?

Note!

$Df(x)=\begin{bmatrix}-\sin x\\ \cos x\end{bmatrix}\neq 0$ for all x

Theorem (MVT)

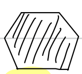
$U\subseteq\mathbb{R}^n$ open, $f:U\to\mathbb{R}^m$ diff if $a,b\in U$ s.t. $L(a,b)\subseteq U$. then $\forall x\in\mathbb{R}^m$ $\exists\, c\in L(a,b)$

s.t. $x\cdot(f(b)-f(a))=x\cdot[Df(c)(b-a)]$

Proof Appendix

Def'n A set $A\subseteq\mathbb{R}^n$ is convex if $L(a,b)\subseteq A$ for all $a,b\in A$

Ex) in $\mathbb{R}^2$



Convex

Not Convex

Remark If $U\subseteq\mathbb{R}^n$ is Convex then the Condition $L(a,b)\subseteq U$ in the MVT is redundant

Corollary

$U\subseteq\mathbb{R}^n$ open, convex. $f:U\to\mathbb{R}^m$ differentiable

If $Df(a)=0$ $\forall a\in U$, then f is Constant.

Proof Let $a,b\in U$, let $\{e_1,e_2,\cdots,e_m\}$ be standard basis for $\mathbb{R}^m$

By the MVT, $\forall e_i$ $\exists\, c_i\in L(a,b)$ s.t. $e_i\cdot(f(b)-f(a))=e_i\cdot \overbrace{Df(c_i)}^{0}(b-a)$

$=0$

$\Rightarrow \forall e_i\ \ e_i\cdot(f(b)-f(a))=0$

$\Rightarrow f(b)-f(a)=0$

Week 8 Appendix

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.

□

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. Since f and g are differentiable at a there exist $r_1, r_2 > 0$ such that f is defined on $B_{r_1}(a)$ and g is defined on $B_{r_2}(a)$. By taking $r = \min\{r_1, r_2\}$, we see that $f \cdot g$ is defined on $B_r(a)$. Therefore we are left to prove that

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

where $X = g(a)Df(a) + f(a)Dg(a)$. Well,

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

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,

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

The result follows. □

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 have that f is defined on some open set $B_{r_1}(a)$ and g is defined on some open set $B_{r_2}(f(a))$. By continuity of f, we may shrink r_1, if necessary, so that $f(B_{r_1}(a)) \subseteq B_{r_2}(f(a))$. Therefore $g \circ f$ is defined on $B_{r_1}(a)$. We are then left to show 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 (Mean Value theorem)**.** Let $U \subseteq \mathbb{R}^n$ be open and let $f : U \to \mathbb{R}^m$ be differentiable. If $a, b \in U$ such that $L(a,b) \subseteq U$, then for all $x \in \mathbb{R}^m$ there exists $c \in L(a,b)$ such that

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

Proof. Consider $\varphi : \mathbb{R} \to \mathbb{R}^n$ given by $\varphi(t) = (1-t)a + tb$ so that $L(a,b) \subseteq \varphi(\mathbb{R})$. Moreover, it may be routinely verified (share your proof on Piazza!) that φ is differentiable with $D\varphi(t) = b - a$. Since U is open and $L(a,b) \subseteq U$ there exists $\delta > 0$ such that $\varphi(t) \in U$ for all $t \in (0-\delta, 1+\delta)$. Then, by the chain rule,

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

for all $t \in (0-\delta, 1+\delta)$. Now fix $x \in \mathbb{R}^m$.

Consider $F : (-\delta, 1+\delta) \to \mathbb{R}$ given by $F(t) = x \cdot (f \circ \varphi)(t)$. By the dot product rule, F is differentiable and

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

By the single-variable MVT, there exists $t_0 \in (0,1)$ such that

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

Hence,

$$x \cdot f(\varphi(1)) - x \cdot f(\varphi(0)) = x \cdot Df(\varphi(t_0))(b-a)$$

and so

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

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

$\square$

9.1 Taylor's Formula

Definition: Higher order partial derivatives are defined recursively by

$$\frac{\partial^k f}{\partial x_1 \partial x_2 \cdots \partial x_k} := \frac{\partial}{\partial x_1}\left(\frac{\partial^{k-1} f}{\partial x_2 \cdots \partial x_k}\right),$$

if it exists. We call k the order of the partial derivative. We also use the notation

$$f_{x_k x_{k-1} \cdots x_1} = \frac{\partial^k f}{\partial x_1 \partial x_2 \cdots \partial x_k}.$$

Also note that I am not assuming the x_i's are distinct here.

Definition: Let $f : U \to \mathbb{R}^m$ be a function on an open set $U \subseteq \mathbb{R}^n$. We say $f \in C^k(U, \mathbb{R}^m)$ if all partial derivatives of f of order less than or equal to k exist on U and are continuous on U. If $m = 1$ we write $C^k(U, \mathbb{R}) = C^k(U)$.

<u>Defn</u> $p \in \mathbb{N}$, $\mathcal{U} \subseteq \mathbb{R}^n$ open, $a \in \mathcal{U}$, $f : \mathcal{U} \to \mathbb{R}$

We define the <u>p^{th} total differential</u> of f at a by $D^p f(a) : \mathbb{R}^n \to \mathbb{R}$

$$D^p f(a)(h_1, h_2, \cdots, h_n)$$

$$= \sum_{i_1=1}^{n} \cdots \sum_{i_p=1}^{n} \frac{\partial^p f(a)}{\partial x_{i_1} \cdots \partial x_{i_p}} \overset{\text{连乘}}{(h_{i_1} h_{i_2} \cdots h_{i_p})}$$

Summing all possible different combinations

Provided it exises

ex) $f : \mathbb{R}^n \to \mathbb{R}$, $a \in \mathbb{R}^n$, $h : (h_1, h_2, \cdots, h_n) \in \mathbb{R}^n$

$$D^1 f(a)(h_1, h_2, \cdots, h_n)$$
$$= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i$$
$$= \nabla f(a) \cdot (h_1, h_2, \cdots, h_n)$$
$$= \nabla f(a) h$$

ex) $f : \mathbb{R}^2 \to \mathbb{R}$, $a \in \mathbb{R}^2$, $(h_1, h_2) \in \mathbb{R}^2$

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

If $f \in C^2(\mathbb{R}^2)$ (by A4), $f_{xy}(a) = f_{yx}(a)$

and so we may simplify ☆

<u>Remark</u> If $f \in C^p(\mathcal{U})$ then $D^p f(c)$ exists and Continuous on $\mathcal{U}$

Theorem (Taylor's theorem)

$p \in \mathbb{N}$, $U \subseteq \mathbb{R}^n$ open + convex, $f \in C^p(U)$

For all $x, a \in U$, $\exists\, c \in L(x,a)$ s.t.

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

Remark $\lim_{x \to a} D^p f(c)(x-a) = D^p f(a)(a-a) = 0$

Proof Appendix

Remark: $U \subseteq \mathbb{R}^n$ open + convex, $f \in C^2(U)$, $\nabla f(a) = 0$

For all $x, a \in U$, $\exists\, c \in L(x,a)$ s.t.

$$f(x) = f(a) + \underbrace{D^1 f(a)(x-a)}_{0} + \frac{1}{2} D^2 f(c)(x-a)$$

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

9.2 Optimization

Goal. Find "extreme" value of multivariable scalar functions $f: \mathcal{U} \to \mathbb{R}$, $\mathcal{U} \subseteq \mathbb{R}^n$ open

Def'n $\mathcal{U} \subseteq \mathbb{R}^n$ open, $a \in \mathcal{U}$ suppose $f: \mathcal{U} \to \mathbb{R}$ is a function

① We call $f(a)$ a local minimum of f if $\exists\, r > 0$ s.t. $f(a) \le f(x)$ $\forall x \in B_r(a)$

② local maximum $(f(a) \ge f(x))$

③ We say $f(a)$ is a local extremum of f if it is either a local min/max of f

Theorem $f: \mathcal{U} \to \mathbb{R}$, $\mathcal{U} \subseteq \mathbb{R}^n$ open if f is diff at $a \in \mathcal{U}$ and $f(a)$ is a local extremum of f, then $\nabla f(a) = 0$

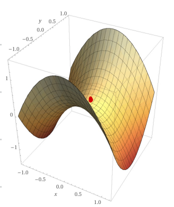
Proof Fix $1 \le i \le n$, suppose $a = (a_1, a_2, \cdots, a_n) \in \mathcal{U}$ and consider $g(x) = f(a_1, \cdots, a_{i-1}, x, a_{i+1}, \cdots, a_n)$

$\therefore$ g has a local extremum at $x = a_i$

$\Longrightarrow$ $g'(a_i) = \frac{\partial f}{\partial x_i}(a) = 0$

$\therefore$ $\forall\, 1 \le i \le n$ $\frac{\partial f}{\partial x_i}(a) = 0 \implies \nabla f(a) = 0$

ex) Warning! $f(x,y) = y^2 - x^2$, $\nabla f(0,0) = (0,0)$ Critical points are only candidates!



$f(0,0)$ is not a local extreme of f

This is something called a saddle point

Def'n $f: \mathcal{U} \to \mathbb{R}$, $\mathcal{U} \subseteq \mathbb{R}^n$ open, suppose f is diff at $a \in \mathcal{U}$. We call a is a saddle point of f

If $\nabla f(a) = 0$ and $\exists\, r > 0$ s.t. for all $0 < \varepsilon < r$ $\exists\, x, y \in B_\varepsilon(a)$ s.t. $f(x) < f(a) < f(y)$

Main Tool:

$\mathcal{U} \subseteq \mathbb{R}^n$ open, $f: \mathcal{U} \to \mathbb{R}$, $a \in \mathcal{U}$, Suppose $f \in C^2(\mathcal{U})$

Consider : $F(f,a) : \mathbb{R}^n \to \mathbb{R}$

$$\boxed{F(f,a)(h_1, h_2, \cdots, h_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\, h_i h_j}$$

i.e $F(f,a) = D^2 f(a)$

ex) $f: \mathbb{R}^2 \to \mathbb{R}$, $f \in C^2(\mathbb{R}^2)$

$$F(f,a)(h_1, h_2) = f_{xx}(a) h_1^2 + f_{yy}(a) h^2 + 2 f_{xy}(a) h_1 h_2$$

Theorem [Second derivative test]

$\mathcal{U} \subseteq \mathbb{R}^n$ open, $f: \mathcal{U} \to \mathbb{R}$, $a \in \mathcal{U}$, $\nabla f(a) = 0$, $f \in C^2(\mathcal{U})$

① If $F(f,a)(n) > 0$, $\forall n \neq 0$ then $f(a)$ is a local min of f

② If $F(f,a)(n) < 0$, $\forall n \neq 0$ then $f(a)$ is a local max of f

③ If $\exists h, k \in \mathbb{R}^n$, s.t. $F(f,a)(h) > 0$ and $F(f,a)(k) < 0$, then a is a saddle point of f

Proof: Appendix

9.3 Examples

Find and classify all extreme values of ...

ex) $f(x,y) = x^4+y^4-4xy+2$, Note $f \in C^2(\mathbb{R}^2)$

Step ① $\nabla f(x,y) = (4x^3-4y,\ 4y^3-4x)$

$\therefore\ \nabla f(x,y) = (0,0)$ iff $x^3=y$ and $y^3=x$

$\iff (x,y) = \{\overset{a}{(0,0)},\ \overset{b}{(1,1)},\ \overset{c}{(-1,-1)}\}$

Step ② $f_{xx}(x,y) = 12x^2$

$f_{yy}(x,y) = 12y^2$

$f_{xy}(x,y) = -4$

$(h_1,h_2) \in \mathbb{R}^2$
$\neq 0$

Step ③ $F(f,a)(h_1,h_2)$

$= f_{xx}(a)h_1^2 + f_{yy}(a)h_2^2 + 2f_{xy}(a)h_1h_2$

$= \underbrace{-8h_1h_2}_{+\text{ or }-}$

$\Rightarrow a$ is a saddle Point

$F(f,b)(h_1,h_2)$

$= f_{xx}(b)h_1^2 + f_{yy}(b)h_2^2 + 2f_{xy}(b)h_1h_2$

$= 4(3h_1^2+3h_2^2-2h_1h_2)$

$\geq 4(h_1^2+h_2^2-2h_1h_2)$

$= 4(h_1-h_2)^2 \geq 0$

$\Rightarrow f(1,1) = 1^4+1^4-4\cdot1\cdot1+2 = 0$ is a local min

$F(f,c)(h_1,h_2)$

$= f_{xx}(c)h_1^2 + f_{yy}(c)h_2^2 + 2f_{xy}(c)h_1h_2$

$= 12h_1^2 + 12h_2^2 - 8h_1h_2$

$\geq 4(h_1^2+h_2^2-2h_1h_2)$

$= 4(h_1-h_2)^2 \geq 0$

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

ex) $\underset{\text{Compact}}{A} = \{(x,y)\in\mathbb{R}^2 : |x|\leq 1,\ |y|\leq 1\}$, $f: A \to \mathbb{R}$, $f(x,y) = \overset{\text{Continuous}}{x^2+y^2+x^2y+4}$

By the EVT. max $f(A)$, min $f(A)$ exists. Find them!

The max/min occur at a) A local extremma
or
b) on the boundary of A

① Checking Local extrema $U = \text{Int}\,A = \{(x,y)\in\mathbb{R}^2 : |x|<1,\ |y|<1\}$, Note $f: U\to\mathbb{R}$ in $C^1(\mathbb{R}^2)$

$\nabla f(x,y) = (2x+2xy,\ 2y+x^2) = (0,0) \iff (x,y) \in \{(0,0),\ \underset{\notin A}{(\sqrt{2},-1)},\ \underset{\notin A}{(-\sqrt{2},-1)}\}$, yet $f(0,0)=4$

② Checking on the boundary $\partial(A) = \{(x,y)\in\mathbb{R}^2 : \max\{|x|,|y|\}=1\}$

$f(1,y) = y^2+y+5 \quad |y|\leq 1$

$f(-1,y) = y^2+y+5 \quad |y|\leq 1$

$f(x,1) = 2x^2+5 \quad |x|\leq 1$

$f(x,-1) = 5 \quad |x|\leq 1$

Of all these, the largest value obtained from one of the above is 7, the smallest is $5-\frac{1}{4} = \frac{19}{4} \to \ \geq f(0,0)$

$\therefore$ min $f(A) = 4$

max $f(A) = 7$

Week 9 Appendix

Theorem. (Taylor's Formula) Let $p \in \mathbb{N}$, $U \subseteq \mathbb{R}^n$ be open and convex, and $f \in C^p(U)$. For all $x, a \in U$ 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 total differential! 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 Formula,

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

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 $c = a + th$.

$\square$

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

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

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 $F(f,a)$ is continuous and positive on K. By the EVT, there exists $m > 0$ such that $m = \min\{F(f,a)(x) : x \in K\}$. For $0 \neq x \in \mathbb{R}^n$ we then see that $\frac{x}{||x||} \in K$ and so

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

□

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

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

for $||h||$ sufficiently small.

Proof. Consider

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

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 Formula that

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

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

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

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 $F(f,a)(h) > 0$ for all $0 \neq h \in \mathbb{R}^n$ then $f(a)$ is a local minimum of f.
2. If $F(f,a)(h) < 0$ 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 $F(f,a)(h) > 0$ and $F(f,a)(k) < 0$ 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

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

for $||h||$ sufficiently small.

1. Suppose $F(f,a)(h) > 0$ for all $0 \neq h \in \mathbb{R}^n$. Let $m > 0$ such that

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

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

   $$f(a+h) - f(a) = \frac{1}{2}F(f,a)(h) + ||h||^2\varepsilon(h) \geq \left(\frac{m}{2} + \varepsilon(h)\right)||h||^2 > 0$$

   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 $t \in \mathbb{R}$,

   $$\begin{aligned} f(a+th) - f(a) &= \frac{1}{2}F(f,a)(th) + ||th||^2\varepsilon(th) \\ &= t^2\left(\frac{1}{2}F(f,a)(h) + ||h||^2\varepsilon(th)\right). \end{aligned}$$

   Letting $t \to 0$, we have that $\varepsilon(th) \to 0$ and so $f(a+th) - f(a)$ takes on the same sign as $F(f,a)(h)$, which can be both positive and negative. Therefore a is a saddle point.

□

10.1 Inverses Function Theorem

Recall: $I=(a,b)$. If $f: I \to R$ is Continuous and injective and $y \in f(I)$ is ① f is diff at $x = f^{-1}(y) \in I$

② $f'(x) \neq 0$

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

Goal: Develop a multivariable version of this theorem

To Generallize the idea of $\frac{1}{f'(x)} = (f'(x))^{-1}$ to something like $Df(x)^{-1}$:

matrix inverse

Def'n $U \subseteq \mathbb{R}^n$, $f: U \to \mathbb{R}^n$ we define the jacobian of f at $a \in U$ by $Jf(a) := \det(Df(a))$

Theorem [Inverse function theorem]

$U \subseteq \mathbb{R}^n$ open. $f \in C'(U, \mathbb{R}^n)$, If $a \in U$ s.t. $Jf(a) \neq 0$

Then $\exists$ open $a \in W \subseteq U$ s.t. (1) f is injective on W

(2) $f^{-1} \in C'(f(W), \mathbb{R}^n)$

(3) For all $y \in f(W)$ $D(f^{-1})(y) = [Df(x)]^{-1}$ where $x = f^{-1}(y)$

Proof: Appendix

ex) $f: \mathbb{R}^2 \to \mathbb{R}^2$, $f(x,y) = (x+y, \sin x + \cos y)$ Note that $f_x(x,y) = (1, \cos x)$, $f_y(x,y) = (1, -\sin y)$ So that $f \in C^1(\mathbb{R}^2, \mathbb{R}^2)$

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

Note. $f(x,y) = (0,1)$ $\iff$ $(x+y, \sin x + \cos y) = (0,1)$

Goal: Preimage

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

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

Case ①: $a = (2k\pi, -2k\pi)$, $k \in \mathbb{Z}$, $J f(a) = \begin{vmatrix} 1 & 1 \\ -1 & 0 \end{vmatrix} = -1 \neq 0$, and so by Inverse function Theorem $\exists$ open $a \in W \subseteq \mathbb{R}^2$

s.t. f is injective on W and $f^{-1} \in C^1(f(W), \mathbb{R}^2)$. Note $(0,1) \in f(W)$

Moreover $[Df(a)]^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$

Case ②: $a = (\frac{\pi}{2} + 2k\pi, -\frac{\pi}{2} - 2k\pi)$, $J f(a) = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1 \neq 0$

Again, $\exists$ open $a \in W$ s.t. $f^{-1} \in C^1(f(W), \mathbb{R}^2)$ with $D(f^{-1})(0,1) = Df(a)^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$

Remark: The way we restrict f to make it injective depends on our choice for $f^{-1}(y)$

10.2 implicit Function Theorem

When/Where can $f(x,y,z)=0$ be solved to express z as a function of x,y? $\{(x,y,z)\in\mathbb{R}^3 : f(x,y,z)=0\} = \{(x,y,g(x,y)) : f(x,y,g(x,y))=0\}$

ex) $f(x,y,z) = x^2+y^2+z^2-1=0$, $\mathcal{U}=\{(x,y,z)\in\mathbb{R}^3 . z>0\}$ open on $\mathcal{U}$, $z=\underbrace{\sqrt{1-x^2-y^2}}_{g(x,y)}$ and $f(x,y,g(x,y))=0$

Theorem [Implicit Function Theorem]

$\mathcal{U}\subseteq\mathbb{R}^{n+p}$ open, $f=(f_1,f_2,\cdots,f_n)\in C^1(\mathcal{U},\mathbb{R}^n)$

Let $x_0\in\mathbb{R}^n$, $t_0\in\mathbb{R}^p$ s.t. $f(x_0,t_0)=0$.

If $\det\left[\frac{\partial f_i}{\partial x_j}(x_0,t_0)\right]_{n\times n}\neq 0$ then $\exists$ open $t_0\in V\subseteq\mathbb{R}^p$ and a unique $g\in C^1(V,\mathbb{R}^n)$ s.t. ① $g(t_0)=x_0$ and
② $f(g(t),t)=0$ $\forall t\in V$

Summary $t\in V\subseteq\mathbb{R}^p\to$ Variables to keep $g(t)\in\mathbb{R}^n\to$ Variables replaced by an implicit of t

ex) $xyz+\sin(x+y+z)=0$, Consider $f(x,y,z)=xyz+\sin(x+y+z)$ so that $f\in C^1(\mathbb{R}^3)$. Note: $f(0,0,0)=0$

$p=2,\ n=1$

Now, $f_z(x,y,z)=xy+\cos(x+y+z)\Rightarrow f_z(\underbrace{0,0}_{t_0\ \in\mathbb{R}^2},\underbrace{0}_{x_0\ \in\mathbb{R}})=1\neq 0$ $\therefore \det[1]=1\neq 0$

By implicit function theorem $\exists$ open $V\subseteq\mathbb{R}^2$ with $\overset{t_0}{(0,0)}\in V$ and $g(x,y)$ in $C^1(V)$

s.t. $\underset{t_0\ \ x_0}{g(0,0)=0}$ and $f(x,y,g(x,y))=0$ for all $(x,y)\in V$. i.e. $z=g(x,y)$ on V

ex) Prove $\exists\ u,v:\mathbb{R}^4\to\mathbb{R}$ 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: $f:\mathbb{R}^6\to\mathbb{R}^2$, $f(\underbrace{u,v}_{\text{replace}},\underbrace{x,y,z,w}_{\text{keep}})=(u^2+v^2+w^2-29,\ \frac{u^2}{x^2}+\frac{v^2}{y^2}+\frac{w^2}{z^2}-17)$ $n=2$, $p=4$ $\Longrightarrow f(\underbrace{4,3}_{x_0},\underbrace{2,1,-1,-2}_{t_0})=0$

and $\det\begin{bmatrix}\frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v}\\ \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v}\end{bmatrix}_{2\times 2}=\begin{vmatrix}2u & 2v\\ \frac{2u}{x^2} & \frac{2v}{y^2}\end{vmatrix}=4uv\left(\frac{1}{y^2}-\frac{1}{x^2}\right)$. This is non-zero at $(4,3,2,1,-1,-2)$

By the implicit function theorem, $\exists$ open $(2,1,-1,-2)\in\mathcal{U}$ and $g\in C^1(\mathcal{U},\mathbb{R}^2)$

s.t. $g(2,1,-1,-2)=(4,3)$ and $\forall (x,y,z,w)\in\mathcal{U}$, $f(g(x,y,z,w),x,y,z,w)=0$

$g=(g_1,g_2)$, $u(x,y,z,w)=g_1(x,y,z,w)$
$v(x,y,z,w)=g_2(x,y,z,w)$ $\Longrightarrow$ $u,v\in C^1(\mathcal{U})$

$u(2,1,-1,-2)=4$, $v(2,1,-1,-2)=3$, $f(g(x,y,z,w),x,y,z,w)=0\Rightarrow f(\underbrace{u(x,y,z,w)}_{u},\underbrace{v(x,y,z,w)}_{v},x,y,z,w)=0$

$\Rightarrow u^2+v^2+w^2=29$
$\frac{u^2}{x^2}+\frac{v^2}{y^2}+\frac{w^2}{z^2}=17$

Week 10 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 $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$ 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 $c_1, c_2, \ldots, c_n \in W$. 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 $f = (f_1, f_2, \ldots, f_n) \in C^1(U, \mathbb{R}^n)$. 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

$$F(x, t) := (f(x, t), t) = (f_1(x, t), \ldots, f_n(x, t), t_1, t_2, \ldots, t_p).$$

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 $t \in V$.

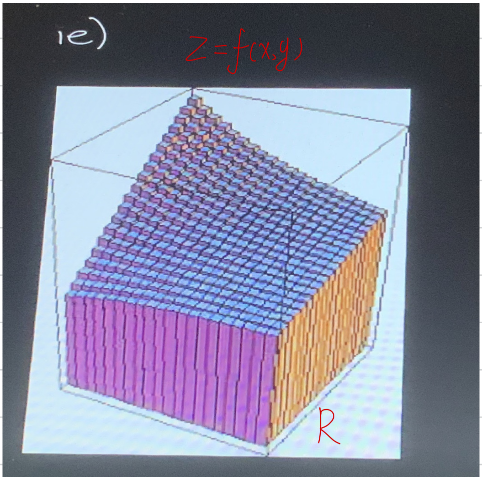
Uniqueness follows from the injectivity of F. (Please share the details on Piazza!)

$\square$

11.1 Jordan Regions 1

In Single variable integration we approximate areas under curves using partitions of intervals and rectangles.

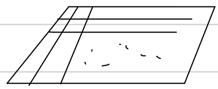
In multivariable integration we approximate "Volumes" of regions under Surfaces using rectangular grids finely covering Jordan regions and "rectangular prisms"



So what are Jordan Regions?

① Regions whose "Volumes" Can be nicely approximated by "rectangles"

② The "nice" regions we integrate over

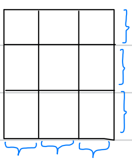


Def'n ① A rectangle in $\mathbb{R}^n$: $R=[a_1,b_1]\times\cdots\times[a_n,b_n]$ with $a_i<b_i$

② The Volume of R: $|R|=(b_1-a_1)(b_2-a_2)\cdots(b_n-a_n)$

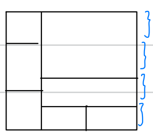
③ A gird on R: A colleetion of rectangles $G=\{R_1, R_2, \cdots R_n\}$ which partition R and made by subdividing the sides of R

Grid



$\mathbb{R}^2$

Not a Gird



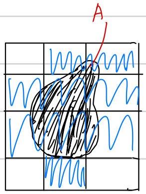
Def'n $A \subseteq \mathbb{R}^n$ bounded.

Let R be a rectangle with $A \subseteq R$ and let $G = \{R_1, \cdots, R_\ell\}$ be a gird on R

We define the outer sum relative to G by $V(A,G) = \sum_{R_i \cap \bar{A} \neq \emptyset} |R_i|$

We then define the Volume of A by: $Vol(A) := \inf \{V(A,G) : G \text{ gird}\}$

eg)



Fact: The definition of $Vol(A)$ is independent of choice of R

ex) $A = ([0,1] \cap \mathbb{Q}) \times ([0,1] \cap \mathbb{Q})$

$R = [0,1] \times [0,1]$, for any gird G

$V(A,G) = \sum_{R_i \cap \bar{A}} |R_i| = \sum_{R_i \cap R} |R_i| = \sum_{R_i} |R_i| = |R| = 1$

$\therefore Vol(A) = 1$ which seems wrong

What happened?

☆ We defined $V(A,G)$ using $\bar{A}$

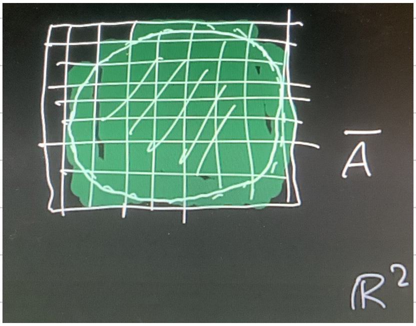
☆ In our example $\bar{A}$ was way "bigger" then A

Idea A (as in the ex) is not soitable to integrable over

Recall The boundary of $A \subseteq \mathbb{R}^n$: $\partial(A) = \bar{A} \setminus Int(A)$

Def'n $A \subseteq \mathbb{R}^n$ bounded, we call A a Jordan Region if $Vol(\partial(A)) = 0$

Idea Jordan regions can be covered well by grids and their volumes are meaningful



11.2 Jordan Regions 2

Properties of Jordan regions:

Prop: Let $\mathbb{R} \subseteq \mathbb{R}^n$ be a rectangle. Then R is a Jordan region with $Vol(R) = |R|$

Proof: $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$, $R \subseteq R$

Let $\varepsilon > 0$ be given. for $\delta > 0$ Consider $R_\delta = [a_1+\delta, b_1-\delta] \times \cdots \times [a_n-\delta, b_n-\delta]$

Then, $|R_\delta| = (b_1 - a_1 - 2\delta) \cdots (b_n - a_n - 2\delta)$

As $\delta \to 0$, $|R_\delta| \to |R|$. So take $\delta > 0$ s.t. $0 < |R| - |R_\delta| < \varepsilon$

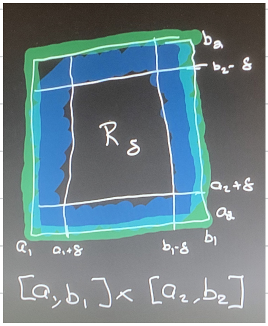
Now Partition $[a_i, b_i]$ by $a_i, a_i+\delta, b_i-\delta, b_i$ and at G be the resulting gird

Say $G = \{H_1, H_2, \cdots, H_\ell\}$ by Construction.

$H_i \cap \partial(R) \neq \emptyset$ iff $H_i \neq R_\delta$

$\therefore V(\partial R, G) = \sum_{H_i \neq R_\delta} |H_i| = |R| - |R_\delta| < \varepsilon \Rightarrow Vol(\partial(R)) = 0$

Let $G = \{R_1, \cdots, R_s\}$ be any grid on R. then $V(R, G) = \sum_{R_i} |R_i| = |R| \Rightarrow Vol(R) = |R|$



Piazza

Prop $A \subseteq B \subseteq \mathbb{R}^n$ bounded $\Rightarrow Vol(A) \leq Vol(B)$

Prop $A \subseteq \mathbb{R}^n$ bounded, $Vol(A) = Vol(\bar{A})$

Prop $A \subseteq \mathbb{R}^n$ bounded, $Vol(A) = 0$. If $B \subseteq A$ then B is a Jordan and $Vol(B) = 0$

Proof $\partial B \subseteq \bar{B} \subseteq \bar{A} \Rightarrow Vol(\partial B) \leq Vol(\bar{A}) = Vol(A) = 0$

$\Rightarrow Vol(\partial B) = 0$

$B \subseteq A$

$\Rightarrow Vol(B) \leq Vol(A)$

Lemma $A \subseteq \mathbb{R}^n$ bounded

Then $Vol(\partial A)=0$ iff $\forall \varepsilon>0$, $\exists$ a finite set of cubes (rectangles with $b_i-a_i=b_j-a_j$) $\{C_k\}_{k=1}^{\ell}$ of some size

s.t. $\bar{A} \subseteq \bigcup_{k=1}^{\ell} C_k$ and $\sum_{k=1}^{\ell} |C_k| < \varepsilon$

Proof: idea $A \subseteq R$

☆ $Vol(\partial A)=0 \iff \forall \varepsilon>0\ \exists\ G$ grid on R s.t. $V(\partial A, G)<\varepsilon$

$\iff \exists$ rectangles $\{B_1, B_2, \cdots, B_p\}$, $\bar{A} \subseteq \bigcup_{k=1}^{p} B_k$ $\sum |B_k| < \varepsilon$

☆ every finite set of rectangles can be arbitrary closely covered by cubes of some size

Prop If $A, B \subseteq \mathbb{R}^n$ are Jordan regions then $A \cup B$ is a Jordan Region with $Vol(A\cup B) \le Vol(A) + Vol(B)$

Proof: Claim $A\cup B$ is a Jordan region. We have that $Vol(\partial A) = Vol(\partial B) = 0$

Let $\varepsilon>0$ be given, by lemma, we may find cubes $\{C_1, \cdots, C_p\}$, $\{D_1, \cdots, D_t\}$ all of the same size

$\partial A \subseteq \bigcup C_i$ $\partial B \subseteq \bigcup D_j$, $\sum |C_i| < \frac{\varepsilon}{2}$ $\sum |D_j| < \frac{\varepsilon}{2}$

If $\{C_1, \cdots, C_p, D_1, \cdots, D_t\}$ is a collection of cubes (same size)

s.t. $\partial(A\cup B) \subseteq \partial(A) \cup \partial(B)$

$\subseteq (\bigcup C_i) \cup (\bigcup D_j)$ and $\sum |C_i| + \sum |D_j| < \varepsilon$

By the lemma, $A\cup B$ is a Jordan Region

Claim: $Vol(A\cup B) \le Vol(A) \cup Vol(B)$

$A\cup B \subseteq R$ ← rectangle

G grid on R

$$V(A\cup B, G) = \sum_{R_i \cap (A\cup B) \neq \emptyset} |R_i| = \sum_{R_i \cap (\overline{A\cup B}) \neq \emptyset} |R_i|$$

$$= \sum_{R_i \cap (\bar{A}) \cup (R_i \cap \bar{B}) \neq \emptyset} |R_i|$$

$$\le \sum_{\substack{R_i \cap \bar{A} \\ \neq \emptyset}} |R_i| + \sum_{\substack{R_i \cap \bar{B} \\ \neq \emptyset}} |R_i|$$

$$= V(A, G) + V(B, G)$$

Take inf's both side: $Vol(A\cup B) \le Vol(A) + Vol(B)$

11.3 Integration 1

Riemann Integration

Recall: $f:[a,b]\to\mathbb{R}$ is integrable iff $U\int_a^b f = \sup\{L(f,P): P \text{ is a Partition}\} = \inf\{U(f,P): P \text{ is a Partition}\} = L\int_a^b f(x)$

where $L(f,P)$ and $U(f,P)$ are the lower and upper sums of f over P

i.e. If $P = a < x_0 < x_1 \cdots < x_n = b$

Then $U(f,P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1})$ $\quad L(f,P) = \sum_{i=1}^{n} m_i(x_i - x_{i-1})$

with $M_i = \sup\{f(x): x\in[x_{i-1}, x_i]\}$ $\quad m_i = \inf\{f(x): x\in[x_{i-1},x_i]\}$

Defn's $A\subseteq\mathbb{R}^n$ Jordan Region

Let $f: A\to R$ be bounded. Let R be a rectangle with $A\subseteq R$ and let $G=\{R_1, R_2, \cdots R_p\}$ be a gird on R

We extrend $f:\mathbb{R}^n\to R$, by setting $f(x)=0 \ \forall x\notin A$

① Upper Sum of f on A wrt G

$U(f,G) = \sum_{R_i\cap A\neq\emptyset} M_i|R_i|$, $M_i = \sup\{f(x): x\in R_i\}$

② Lower Sum of f on A wrt G

$L(f,G) = \sum_{R_i\cap A\neq\emptyset} m_i|R_i|$, $m_i = \inf\{f(x): x\in R_i\}$

③ upper integral of f on A:

$U\int_A f(x)dx = \inf\{U(f,G): G \text{ is a gird}\}$

③ Lower integral of f on A:

$L\int_A f(x)dx = \sup\{L(f,G): G \text{ is a gird}\}$

Fact The defn of $U\int_A f(x)dx$, $L\int_A f(x)dx$ exists and do not depend on choice of R

Def'n $A\subseteq\mathbb{R}^n$ Jordan Region, A bounded function $f: A\to\mathbb{R}$ is said to be (Riemann) integrable on A iff $\forall\varepsilon>0$, $\exists$ grid G s.t. $\underbrace{U(f,G)-L(f,G)}_{\geqslant 0}<\varepsilon$

Prop $A \subseteq \mathbb{R}^n$ a Jordan Region, $f: A \to \mathbb{R}$ bounded, TFAE

① f is integrable on A

② $L\int_A f(x)dx = U\int_A f(x)dx := \int_A f(x)dx$

Proof:

① ⇒ ②: Suppose f is integrable on A

Let $\varepsilon > 0$ be given, so $\exists$ grid G s.t. $U(f,G) - L(f,G) < \varepsilon$

$$\Longrightarrow 0 \le U\int_A f(x)dx - L\int_A f(x)dx$$

$$\le U(f,G) - L(f,G) < \varepsilon$$

$$\Longrightarrow U\int_A f(x)dx = L\int_A f(x)dx$$

② ⇒ ①: Assume $U\int_A f(x)dx = L\int_A f(x)dx$

Let $\varepsilon > 0$ be given. We may find grids G_1, G_2

$U(f,G_1) \le U\int_A f(x)dx + \frac{\varepsilon}{2}$ $\qquad$ $L\int f(x)dx - \frac{\varepsilon}{2} \le L(f,G_2)$

Consider $G = G_1 \cup G_2$. Note $U(f,G) \le U(f,G_1)$, $L(f,G) \ge L(f,G_2)$ } exercise

$$\therefore\ U(f,G) - L(f,G) \le U(f,G_1) - L(f,G_2)$$

$$\le \cancel{U\int_A f(x)dx} + \frac{\varepsilon}{2} - \cancel{L\int_A f(x)dx} + \frac{\varepsilon}{2}$$

$$\le \varepsilon$$

Prop: $A \subseteq \mathbb{R}^n$ a Jordan Region, $f: A \to \mathbb{R}$ bounded

Let $A \subseteq R$ be a rectangle. then $f: A \to \mathbb{R}$ is integrable iff $g: R \to \mathbb{R}$ given by $g(x) = \begin{cases} f(x) & x \in A \\ 0 & x \in R \setminus A \end{cases}$ is integrable

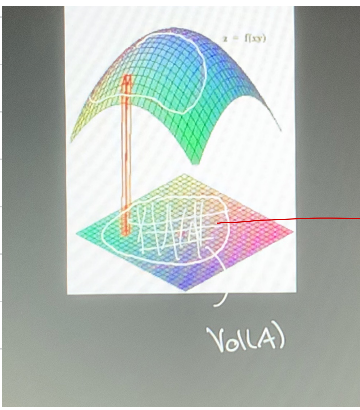
Moreover, $\int_A f(x)dx = \int_R g(x)dx$

why? $\forall$ grids G on R, $L(f,G) = L(g,G)$, $U(f,G) = U(g,G)$

Big Picture:

① If $A \subseteq \mathbb{R}^n$ is a Jordan Region, then A can be covered "nicely" by rectangles (in $\mathbb{R}^n$)

② If $z = f(x_1, x_2, \cdots, x_n)$ is integrable on A, then we can approximate the volume between $z = f(x_1, x_2, \cdots, x_n)$ and $z = 0$ over A using rectangles (in $\mathbb{R}^n$)



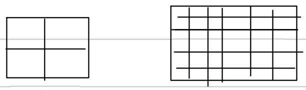
Jordan Region A

11.4 Integration 2

Question: Why are the Jordan Regions the "right" sets to integrable over?

Def'n Let G be a grid on a rectangle R, we say a grid G' on R is <u>finer</u> than G if G' can be obtained from G by partitioning the sides of R even further

ex)



G $\qquad$ G'

Remark. $U(f, G') \le U(f, G)$, $L(f, G') \ge L(f, G)$

Theorem $A \subseteq \mathbb{R}^n$ Jordan Region, $f: A \to \mathbb{R}$ bounded. $\forall \varepsilon > 0$ $\exists G_0$ s.t. if $G = \{R_1, \cdots, R_p\}$ is a grid finer than G_0, then $\left| U\int_A f(x)dx - \sum_{R_i \subseteq Int(A)} M_i |R_i| \right| < \varepsilon$, $\left| L\int_A f(x)dx - \sum_{R_i \subseteq Int(A)} m_i |R_i| \right| < \varepsilon$ where m_i, M_i are as before

Proof: Let $\varepsilon > 0$ be given. Let $M > 0$ s.t. $|f(x)| \le M$ for all $x \in A$

Since $Vol(\partial A) = 0$, we may find a grid G_1 s.t. $V(\partial A, G_1) < \frac{\varepsilon}{2M}$

Moreover, we may find a grid G_2 s.t. $U\int_A f(x)dx \le U(f, G_2) \le U\int_A f(x)dx + \frac{\varepsilon}{2}$

Set $G_0 := G_1 \cup G_2$, Let G be a gird which is finer than G_0

Then, $U\int_A f(x)dx \le U(f, G) \le U(f, G_0) \le U\int_A f(x)dx + \frac{\varepsilon}{2}$

$\Rightarrow$ Say $G = \{R_1, R_2, \cdots, R_p\}$ then

$$\begin{aligned}
\left| U\int_A f(x)dx - \sum_{R_i \subseteq Int(A)} M_i |R_i| \right| &\le \frac{\varepsilon}{2} + \left| U(f, G) - \sum_{R_i \subseteq Int(A)} M_i |R_i| \right| \\
&\le \frac{\varepsilon}{2} + \left| \sum_{\substack{R_i \cap A \\ \neq \emptyset}} M_i |R_i| - \sum_{R_i \subseteq Int(A)} M_i |R_i| \right| \\
&= \frac{\varepsilon}{2} + \left| \sum_{\substack{R_i \cap \partial A \\ \neq \emptyset}} M_i |R_i| \right| \\
&\le \frac{\varepsilon}{2} + \sum_{\substack{R_i \cap \partial A \\ \neq \emptyset}} |M_i| \cdot |R_i| \\
&\le \frac{\varepsilon}{2} + \sum_{\substack{R_i \cap \partial A \\ \neq \emptyset}} M \cdot |R_i| \\
&= \frac{\varepsilon}{2} + M \sum_{\substack{R_i \cap \partial A \\ \neq \emptyset}} |R_i| \\
&= \frac{\varepsilon}{2} + M \cdot V(\partial A, G) \\
&< \frac{\varepsilon}{2} + M \cdot \frac{\varepsilon}{2M} \\
&= \varepsilon
\end{aligned}$$

The Proof using lower integral is similarly as the upper integral.

Theorem $A \subseteq \mathbb{R}^n$ a closed bounded Jordan Region. If $f: A \to \mathbb{R}$ is continuous then f is integrable

Proof Suppose $f: A \to \mathbb{R}$ is continuous, Since A is compact, f is bounded and f is uniformly continuous (Compact + Continuous $\Rightarrow$ Uniformly Continuous)

Let $\varepsilon > 0$ be given by previous results, $\exists\, G_0$ s.t. if G finer than G_0 (say $G = \{R_1, \cdots, R_k\}$)

Then, $\left| U\int_A f(x)dx - \sum_{R_i \subseteq Int(A)} M_i |R_i| \right| < \frac{\varepsilon}{2}$, $\left| L\int_A f(x)dx - \sum_{R_i \subseteq Int(A)} m_i |R_i| \right| < \frac{\varepsilon}{2}$

$\Longrightarrow \left| U\int_A f(x)dx - L\int_A f(x)dx - \sum_{R_i \subseteq Int(A)} (M_i - m_i)|R_i| \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$

Since f is uniformly continuous, $\exists\, \delta > 0$ s.t. If $x, y \in A$ with $\|x-y\| < \delta$ then $\|f(x) - f(y)\| < \varepsilon$.

Take a grid $G = \{R_1, \cdots, R_p\}$, if $x, y \in R_i$, then $\|x-y\| < \delta$

From the above: $\left| U\int_A f(x)dx - L\int_A f(x)dx \right| \le \varepsilon + \left| \sum_{R_i \subseteq Int(A)} (M_i - m_i)|R_i| \right| \le \varepsilon + \sum_{R_i \subseteq Int(A)} |M_i - m_i| \cdot |R_i| \le \varepsilon + \sum_{R_i \subseteq Int(A)} \varepsilon \cdot |R_i| = \varepsilon + \varepsilon \sum_{R_i \subseteq Int(A)} |R_i| \le \varepsilon + \varepsilon |R|$

as $\varepsilon \to 0$, $\varepsilon + \varepsilon|R| \to 0$

$\therefore U\int_A f(x)dx = L\int_A f(x)dx \Rightarrow$ Integrable

Remark [Piazza]: we can drop the assumption that A is closed if we insist that f uniformly continuous

ex) $f(x) = \frac{1}{x}$ is continuous on $(0,1)$ but not integrable.

12.1 Integration 4

Properties of Integrals:

Theorem: $A \subseteq \mathbb{R}^n$ a Jordan Region

$$\int_A 1 \, dx = Vol(A)$$

Why? $A \subseteq R$ rectangle

$G = \{G_1, G_2, \cdots, G_L\}$ gird on R

$f: R \to R$, $f(x)=1$ on A, $f(x)=0$ on $R \setminus A$

Note, f is integrable on A

☆ $U(f,G) = \sum_{R_i \cap A \neq \emptyset} M_i |R_i| \leq \sum_{\substack{R_i \cap \bar{A} \\ \neq \emptyset}} M_i |R_i| \leq \sum_{\substack{R_i \cap \bar{A} \\ \neq \emptyset}} |R_i| = V(A,G)$

Take inf's $\int_A f(x)dx \leq Vol(A)$, $Vol(\partial A) = 0$. $\varepsilon > 0$

Pick G s.t. $V(\partial A, G) < \varepsilon$ ☆ $\int_A 1 dx \geq L(f,G) = \sum_{R_i \cap A \neq \emptyset} m_i |R_i| = \sum_{R_i \in A} |R_i|$

$\geq \sum_{R_i \cap \bar{A}} |R_i| - \sum_{\substack{R_i \cap \partial A \\ \neq \emptyset}} |R_i|$

$= V(A,G) - V(\partial A, G)$

$\geq V(A,G) - \varepsilon$

$\geq Vol(A) - \varepsilon$

$\Longrightarrow \int_A 1 dx \geq Vol(A)$

Prop $A \subseteq \mathbb{R}^n$ Jordan Region . f, g $A \to R$ integrable

① $f+g$ integrable with $\int_A f(x) + g(x) \, dx = \int_A f(x)dx + \int_A g(x)dx$

② If $\alpha \in R$, then αf is integrable with $\int_A \alpha f(x)dx = \alpha \int_A f(x)dx$

Why?

① $\varepsilon > 0$. choose a grid G s.t. $U(f,G) - \varepsilon < \int_A f(x)dx < L(f,G) + \varepsilon$

$U(g,G) - \varepsilon < \int_A g(x)dx < L(g,G) + \varepsilon$

$\Longrightarrow U(f,G) + U(g,G) - 2\varepsilon < \int_A f(x)dx + \int_A g(x)dx < L(f,G) + L(g,G) + 2\varepsilon$

Note $U(f+g,G) \leq U(f,G) + U(g,G)$ and $L(f+g,G) \geq L(f,G) + L(g,G)$

$\Rightarrow U(f+g,G) - 2\varepsilon < \int_A f(x)dx + \int_A g(x)dx < L(f+g,G) + 2\varepsilon$

$\Rightarrow f+g$ integrable

Moreover $U \int_A (f+g)(x)dx - 2\varepsilon < \int_A f(x)dx + \int_A g(x)dx < L \int_A (f+g)(x)dx + 2\varepsilon$

Prop $E \subseteq \mathbb{R}^n$ Jordan Region

If $A, B \subseteq E$ are JR's s.t. $Vol(A\cap B) = 0$ and $f: E \to \mathbb{R}$ is integrable over both A and B, then f is integrable over $A\cup B$

With $\int_{A\cup B} f(x)dx = \int_A f(x)dx + \int_B f(x)dx$

Why!

For $\varepsilon > 0$, $\exists$ grid G_0 s.t. If $G = \{R_1, \dots, R_l\}$ is finer than G_0

then ① $\left|\int_A f(x)dx - \sum_{R_i \subseteq Int(A)} M_i |R_i|\right| < \varepsilon$ ② $\left|\int_B f(x)dx - \sum_{R_i\subseteq Int(A)} M_i|R_i|\right| < \varepsilon$ ③ $\left| U\int_{A\cup B} f(x)dx - \sum_{R_i \subseteq Int(A\cup B)} M_i|R_i|\right| < \varepsilon$

We may also assume each $V(A\cap B, G) < \varepsilon$

$M := \max\{|M_1|, \dots, |M_l|\}$ ☆ $U\int_{A\cup B} f(x)dx \le \varepsilon + \sum_{R_i\subseteq Int(A\cup B)} M_i|R_i| \le \varepsilon + \sum_{R_i\subseteq Int(A)} M_i|R_i| + \sum_{R_i\subseteq Int(B)} M_i|R_i| + \sum_{R_i\cap(A\cap B)\neq\emptyset} M_i|R_i|$

$$\le 3\varepsilon + \int_A f(x)dx + \int_B f(x)dx + M\cdot V(A\cap B, G)$$

$$\le (3+M)\varepsilon + \int_A f(x)dx + \int_B f(x)dx$$

$$\Longrightarrow U\int_{A\cup B} f(x)dx \le \int_A f(x)dx + \int_B f(x)dx$$

$\cdot\,\cdot - -$ $- - - -$ $-$ Hw: $\int_A f(x)dx + \int_B f(x)dx \le L\int_{A\cup B} f(x)dx$

Prop $A \subseteq \mathbb{R}^n$ Jordan Region. $f, g: A \to \mathbb{R}$ bounded

① If $B \subseteq A$ is s.t. $Vol(B) = 0$, then f is integrable on B and $\int_B f(x)dx = 0$

② If f is integrable on A and $B \subseteq A$ s.t. (a) $Vol(B) = 0$ $\Longrightarrow$ g is integrable on A with $\int_A f(x)dx = \int_A g(x)dx$
(b) $f = g$ on $A\setminus B$

Why?

① If $Int(B) \neq \emptyset$ then $\exists$ rectangle $R \subseteq B \Rightarrow \underset{\neq 0}{Vol(R)} < Vol(B)$ Contradiction $\therefore Int(B) = \emptyset$

Let $\varepsilon > 0$ we may find $G = \{R_1, \dots, R_l\}$ s.t. $\left|U\int_B f(x)dx - \sum_{R_i\subseteq Int(B)} M_i|R_i|\right| < \varepsilon$ $\left|L\int_B f(x)dx - \sum_{R_i\subseteq Int(B)} m_i|R_i|\right| < \varepsilon$

② $\int_A f(x)dx = \underbrace{\int_B f(x)dx}_{0} + \int_{A\setminus B} f(x)dx = \underbrace{\int_B g(x)dx}_{0} + \int_{A\setminus B} g(x)dx = \int_A g(x)dx$

Prop $A \subseteq \mathbb{R}^n$ Jordan Region, $f, g: A \to \mathbb{R}$ integrable

① If $f(x) \le g(x)$ $\forall x \in A$, then $\int_A f(x)dx \le \int_A g(x)dx$

② If $m \le f(x) \le M$, $\forall x \in A$ then $m\,Vol(A) \le \int_A f(x)dx \le M \cdot Vol(A)$

③ $|f|$ is integrable on A with $|\int_A f(x)dx| \le \int_A |f(x)|dx$

Why!

① $\forall G$, $L(f, G) \le L(g, G) \Longrightarrow \int_A f(x)dx \le \int_B g(x)dx$

③ $-|f(x)| \le f(x) \le |f(x)|$

Check: $U(|f|, G) - L(|f|, G) \le U(f, G) - L(f, G) < \varepsilon$

12.2 Fubinis Theorem

Q: Can we use multiple single-variable integrals to compute integrals of multivariable function?

Q: Would order matter? i.e. $\int_c^d \int_a^b f(x,y)\,dx\,dy = \int_a^b \int_c^d f(x,y)\,dy\,dx$?

Notation

$B \subseteq \mathbb{R}^2$ Jordan Region, $f: B \to \mathbb{R}$ integrable : $\int_B f(v)\,dv \equiv \iint_B f(x,y)\,dA$

$B \subseteq \mathbb{R}^3$ Jordan Region ; $f: B \to \mathbb{R}$ integrable : $\int_B f(v)\,dv \equiv \iiint_B f(x,y,z)\,dV$

Lemma. $R = [a,b] \times [c,d] \subseteq \mathbb{R}^2$. If $f: R \to \mathbb{R}$ 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 $L\iint_R f(x,y)\,dA \le L\int_a^b \left(\int_c^d f(x,y)\,dy\right) dx \le U\int_a^b \left(\int_c^d f(x,y)\,dy\right) dx \le U\iint_R f(x,y)\,dA$

Proof The middle inequality is trivial, we will prove the last inequality and leave the first for a Piazza contribution.

Let $\varepsilon > 0$ be given, choose a grid G on R s.t. $U(f,G) - \varepsilon \le U\iint_R f(x,y)\,dA$

Say $G = \{R_{ij} : 1 \le i \le k,\ 1 \le j \le l\}$, $R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$ where $x_0 = a$, $x_k = b$, $y_0 = c$, $y_l = d$

Set $M_{ij} = \sup\{f(v) : v \in R_{ij}\}$ $\therefore$

$$\begin{aligned}
U\int_a^b \left(\int_c^d f(x,y)\,dy\right) dx &= \sum_{i=1}^{k} U\int_{x_{i-1}}^{x_i} \left(\sum_{j=1}^{l} \int_{y_{j-1}}^{y_j} f(x,y)\,dy\right) dx \\
&\le \sum_{i=1} \sum_{j} U\int_{x_{i-1}}^{x_i} \left(\int_{y_{j-1}}^{y_j} f(x,y)\,dy\right) dx \qquad \left(\text{why? } U\int_a^b (f+g)(x)\,dx \le U\int_a^b f(x)\,dx + U\int_a^b g(x)\,dx\right) \\
&\le \sum_{i} \sum_{j} \int_{x_{i-1}}^{x_i} \left(\int_{y_{j-1}}^{y_j} M_{ij}\,dy\right) dx \\
&= \sum_{i} \sum_{j} M_{ij}\,(x_i - x_{i-1})(y_j - y_{j-1}) \\
&= \sum_{R_{ij}} M_{ij}\,|R_{ij}| \\
&= U(f,G) \\
&\le U\iint_R f(x,y)\,dA + \varepsilon
\end{aligned}$$

Theorem [Fubini's Theorem]

$R=[a,b]\times[c,d]\subseteq\mathbb{R}^2$, $f: R\to\mathbb{R}$ 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 equals $L\int_a^b\int_c^d f(x,y)\,dy\,dx=U\int_a^b\int_a^d f(x,y)\,dy\,dx$

$\therefore\ \iint_R f(x,y)\,dA=\int_a^b\left(\int_c^d f(x,y)\,dy\right)dx$

Reversing the role of x,y in the lemma proves the theorem

Remark we call $\int_{a_n}^{b_n}\int_{a_{n-1}}^{b_{n-1}}\cdots\int_{a_1}^{b_1} f(x_1,x_2,\cdots,x_n)\,dx_1\cdots dx_n$ Iterated Integral

ex) $R=[1,2]\times[0,\pi]$ $\iint_R y\sin(xy)\,dA$

Note f, $f(x,\cdot)$ $f(\cdot,y)$ are all continuous on closed $\mathbb{R}$'s $\Rightarrow$ integrable

$$\iint_R y\sin(xy)\,dA\overset{FT}{=}\int_0^\pi\int_1^2 y\sin(xy)\,dx\,dy=\int_0^\pi\left[-\cos(xy)\right]_{x=1}^{x=2}dy=\int_0^\pi-\cos(2y)+\cos(y)\,dy=\left[-\frac{1}{2}\sin(2y)+\sin(y)\right]_0^\pi=0$$

12.3 Iterated Integrals

Generalizing Fubin:

Theorem: $R=[a_1,b_1]\times\cdots\times[a_n,b_n]\subseteq\mathbb{R}^n$, $f: R\to\mathbb{R}$ integrable. $R_n=[a_1,b_1]\times\cdots\times[a_{n-1},b_{n-1}]$

If $f(x,\cdot)$ is integrable for all $x\in R_n$ then $\int_{a_n}^{b_n} f(x,t)\,dt$ integrable on R_n

and $\int_R f(v)\,dv = \int_{R_n}\int_{a_n}^{b_n} f(x,t)\,dt\,dx$

Remark: If $f: R\to\mathbb{R}$ is continuous, then $\int_R f(v)\,dv = \int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_n}^{b_n} f(x_1,\dots,x_n)\,dx_n\cdots dx_1$

Goal: use iterated integrals to integrate over "nice" regions which are not rectangles

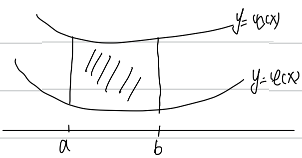
For simplicity we shall work in $\mathbb{R}^2$ and $\mathbb{R}^3$

Def'n 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 $\varphi,\psi: [a,b]\to\mathbb{R}$

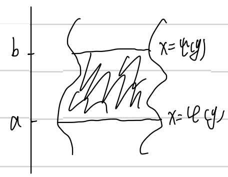
② We say $A\subseteq\mathbb{R}^2$ is type ② if $A=\{(x,y): y\in[a,b],\ \varphi(y)\le x\le\psi(y)\}$ for some continuous $\varphi,\psi: [a,b]\to\mathbb{R}$

eg) In $\mathbb{R}^2$

Type①



Type②



Def'n $A\subseteq\mathbb{R}^3$

① Type 1: $A=\{(x,y,z): (x,y)\in H,\ \varphi(x,y)\le z\le\psi(x,y)\}$

② Type 2 $A=\{(x,y,z): (x,z)\in H,\ \varphi(x,z)\le y\le\psi(x,z)\}$

③ Type 3 $A=\{(x,y,z): (y,z)\in H,\ \varphi(y,z)\le x\le\psi(y,z)\}$

Where $H\subseteq\mathbb{R}^3$ is a closed Jordan region and $\varphi,\psi: A\to\mathbb{R}$ are continuous

Fact Regions of type 1, 2, or 3 are Jordan Regions

Theorem $A \subseteq \mathbb{R}^2$, $f: A \to \mathbb{R}$ 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 $A = \{(x,y) : y \in [a,b],\ \varphi(y) \le x \le \psi(y)\}$ is type 2, 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, $\int_A f(v)\,dv = \int_R f(v)\,dv = \int_a^b \int_c^d f(x,y)\,dy\,dx$. However, $f(x,y) = 0$ if it is not the case that $\varphi(x) \le y \le \psi(x)$

$\therefore \int_A f(v)\,dv = \int_a^b \int_{\varphi(x)}^{\psi(x)} f(x,y)\,dy\,dx$

Theorem $A \subseteq \mathbb{R}^3$, $f: A \to \mathbb{R}$ continuous

① If A is type 1 then $\int_A f(v) = \int_H \int_{\varphi(x,y)}^{\psi(x,y)} f(u,z)\,dz\,du$

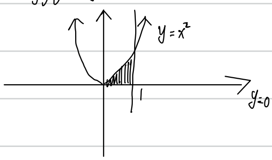
② . — — — etc

③ . - — —

12.4 Examples

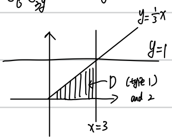
ex) Let $D \subseteq \mathbb{R}^2$ be the region bounded by $y=0,\ y=x^2,\ x=1$

Compute $\iint_D x\cos y\, dA$



$$\iint_D x\cos y\, dA = \int_0^1\int_0^{x^2} x\cos y\, dy\, dx$$
$$= \int_0^1 [x\sin y]_{y=0}^{y=x^2}\, dx$$
$$= \int_0^1 x\sin x^2\, dx$$
$$= \left[-\cos x^2\right]_{x=0}^{x=1}$$
$$= -\frac{1}{2}\cos(1) + \frac{1}{2}$$

ex) $\int_0^1\int_{3y}^3 e^{x^2}\, dx\, dy$



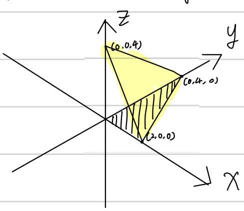
$x=3y \iff y=\frac{1}{3}x$

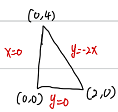
$3y \le x \le 3$

$0 \le y \le 1$

$$\int_0^1\int_{3y}^3 e^{x^2}\, dx\, dy = \iint_D e^{x^2}\, dA = \int_0^3\int_0^{\frac{1}{3}x} e^{x^2}\, dy\, dx = \int_0^3 \left[e^{x^2}y\right]_{y=0}^{y=\frac{1}{3}x} dx$$
$$= \int_0^3 \frac{1}{3}xe^{x^2}\, dx$$
$$= \left[\frac{1}{6}e^{x^2}\right]_0^3$$
$$= \frac{1}{6}(e^9-1)$$

ex) Find the volume of the tetrahedron T enclosed by $x=0,\ y=0,\ z=0$ and $2x+y+z=4$





$T = \{(x,y,z) : 0 \le x \le 2,\ 0 \le y \le -2x+4,\ 0 \le z \le 4-2x-y\}$

$H = \{(x,y) : 0 \le x \le 2,\ 0 \le y \le -2x+4\}$

↳ Type I

$T = \{(x,y,z) : (x,y) \in H,\ 0 \le z \le 4-2x-y\}$

$$\iiint_T 1\, dV = \iint_H \underbrace{\int_0^{4-2x-y} 1\, dz}_{\text{continuous}}\, dA = \int_0^2\int_0^{-2x+4}\int_0^{4-2x-y} 1\, dz\, dy\, dx$$
$$= \int_0^2\int_0^{-2x+4} 4-2x+y\, dy\, dx$$
$$= \int_0^2 \left[(4-2x)y + \frac{1}{2}y^2\right]_{y=0}^{y=-2x+4} dx$$
$$= \int_0^2 (4-2x)^2 - \frac{1}{2}(4-2x)^2$$
$$= \int_0^2 \frac{1}{2}(4-2x)^2\, dx$$
$$= \frac{16}{3}$$

12.5 Change of Variables

Recall $f: A \to \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$, $a \in A$, the Jacobian of f at a: $Jf(a) = \det(Df(a))$

Theorem $U \subseteq \mathbb{R}^n$ open, $A \subseteq U$ 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 $Vol(B) = 0$

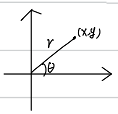
② φ is injective on $A \setminus B$

③ $J\varphi(a) \neq 0 \quad \forall\ a \in A \setminus B$

and suppose $f: \varphi(A) \to \mathbb{R}$ is continuous, then $\varphi(A)$ is a Jordan Region, f is integrable on $\varphi(A)$ and

$$\int_{\varphi(A)} f(x)\,dx = \int_A f(\varphi(x))\,|Jf(x)|\,dx$$

Polar Coordinates



$r = \sqrt{x^2+y^2}$
$x = r\cos\theta$
$y = r\sin\theta \qquad 0 \le \theta \le 2\pi$

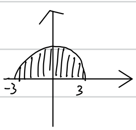
We call (r,θ) the Polar Coordinates of $(x,y) \in \mathbb{R}^2$, Consider $\varphi \in C^1(\mathbb{R}^2, \mathbb{R}^2)$ given by $\varphi(r,\theta) = (r\cos\theta, r\sin\theta)$

Note: φ is injective on $\mathbb{R}^2 \setminus \underbrace{\{(0,\theta) : 0 \le \theta \le 2\pi\}}_{Vol=0}$

$$|J\varphi(r,\theta)| = \left|\det\begin{bmatrix}\cos & -r\sin\theta \\ \sin\theta & r\cos\theta\end{bmatrix}\right| = |r| = r$$

$$\iint_{\varphi(D)} f(x,y)\,dA = \iint_D f(r\cos\theta, r\sin\theta)\,r\,dA$$

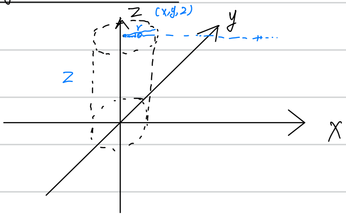
ex) $\iint_D \cos(x^2+y^2)\,dA$, D is the region bounded by $x^2+y^2=9$ and above x-axis



$D = \{(r,\theta) : 0 \le r \le 3,\ 0 \le \theta \le \pi\}$

$$\begin{aligned}\iint_D \cos(x^2+y^2)\,dA &= \iint_D \cos(r^2)\,r\,dA \\ &= \int_0^\pi \int_0^3 \cos(r^2)\,r\,dr\,d\theta \\ &= \int_0^\pi \left[\tfrac{1}{2}\sin(r^2)\right]_0^3 dr \\ &= \int_0^\pi \tfrac{1}{2}\sin(9)\,d\theta = \tfrac{\pi}{2}\sin(9)\end{aligned}$$

Cylindrical Coordinates



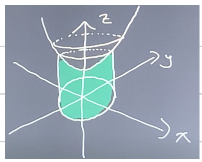
We call (r,θ,z) the cylindrical coordinates of (x,y,z).

$\varphi(r,\theta,z) = (r\cos\theta, r\sin\theta, z)$ $\quad$ $|J\varphi(r,\theta,z)| = \left| \det\begin{bmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \right| = r$ $\quad$ $\iiint_{\varphi(A)} f(x,y,z)\,dV = \iiint_A f(r\cos\theta, r\sin\theta, z)\, r\,dV$

ex) $\iiint_A e^z\,dV$. A enclosed by ① the Paraboloid $z = 1+x^2+y^2$

② the cylinder $x^2+y^2=5$

③ xy-plane



$A = \{(r,\theta,z) : 0 \le r \le \sqrt{5},\ 0 \le \theta \le 2\pi,\ 0 \le z \le 1+\underbrace{x^2+y^2}_{r^2} = 1+r^2\}$

$$\iiint e^z dV = \int_0^{\sqrt{5}}\int_0^{2\pi}\int_0^{1+r^2} e^z \cdot r\,dz\,d\theta\,dr = \int_0^{\sqrt{5}}\int_0^{2\pi} re^{1+r^2} - r\,d\theta\,dr$$

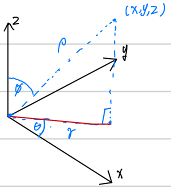
$$= 2\pi\int_0^{\sqrt{5}} re^{1+r^2} - r\,dr$$

$$= 2\pi\left[\tfrac{1}{2}e^{1+r^2} - \tfrac{1}{2}r^2\right]_0^{\sqrt{5}}$$

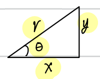
$$= 2\pi\left(\tfrac{1}{2}e^6 - \tfrac{5}{2} - \tfrac{1}{2}e\right)$$

$$= \pi(e^6 - 5 - e)$$

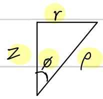
Spherical Coordinates



$0 \le \theta \le 2\pi$
$0 \le \phi \le \pi$



$x = r\cos\theta$
$y = r\sin\theta$



$r = \rho\sin\phi$
$z = \rho\cos\phi$

$x = \rho\sin\phi\cos\theta$, $y = \rho\sin\phi\sin\theta$
$z = \rho\cos\phi$, $x^2+y^2+z^2 = \rho^2$

$\rho^2 s\phi$

Consider $\varphi(\rho,\theta,\phi) = (\rho\sin\phi\cos\theta,\ \rho\sin\phi\sin\theta,\ \rho\cos\phi)$

$$\left|J\varphi(\rho,\theta,\phi)\right| = \left|\det\begin{bmatrix} \sin\phi\cos\theta & -\rho\sin\phi\sin\theta & \rho\cos\phi\cos\theta \\ \sin\phi\sin\theta & \rho\sin\phi\cos\theta & \rho\cos\phi\sin\theta \\ \cos\phi & 0 & -\rho\sin\phi \end{bmatrix}\right| = \rho^2\sin\phi$$

$$\iiint_{\varphi(A)} f(x,y,z)\,dV = \iiint_A f(\rho\sin\phi\cos\theta,\ \rho\sin\phi\sin\theta,\ \rho\cos\phi)\cdot\rho^2\sin\phi\,dV$$

ex) Find the Volume of the sphere $x^2+y^2+z^2=a^2$

$S = \{(\rho,\theta,\phi) : 0 \le \rho \le a,\ 0 \le \theta \le 2\pi,\ 0 \le \phi \le \pi\}$

$$\mathrm{Vol}(S) = \int_S 1\,dV = \iiint_S 1\,dx\,dy\,dz = \int_0^a\int_0^{2\pi}\int_0^{\pi} \rho^2\sin\phi\,d\phi\,d\theta\,d\rho = \int_0^a\int_0^{2\pi}\left[-\rho^2\cos\phi\right]_0^{\pi} d\theta\,d\rho$$

$$= \int_0^a\int_0^{2\pi} 2\rho^2\,d\theta\,d\rho$$

$$= 2\pi\int_0^a 2\rho^2\,d\rho$$

$$= 2\pi\left[\tfrac{2}{3}\rho^3\right]_0^a$$

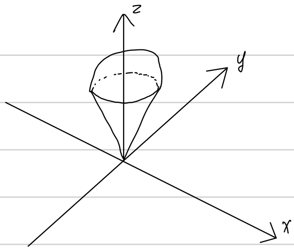
$$= \frac{4\pi^2}{3}a^3$$

ex) Find the volume of the solid which

① lies above the cone $z = \sqrt{x^2+y^2}$

and is

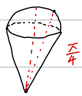
② below the sphere $x^2+y^2+z^2 = z$



$x^2+y^2+z^2 = z \iff x^2+y^2+(z-\frac{1}{2})^2 = \frac{1}{4}$

Cone: $\rho\cos\phi = \sqrt{\rho^2\sin^2\phi\cos^2\theta + \rho^2\sin^2\theta\sin^2\theta}$ $\quad C = \{(\rho,\theta,\phi) : \rho = 0,\ \phi = \frac{\pi}{4}\}$

$= \rho\sin\phi$



Sphere: $\rho^2 = \rho\cos\phi$, $S = \{(\rho,\theta,\phi) : \rho = 0 \text{ or } \rho = \cos\phi\}$

Letting D be the solid $\iiint_D 1\, dV = \int_0^{2\pi}\int_0^{\frac{\pi}{4}}\int_0^{\cos\phi} \rho^2 \sin\phi\, d\rho$

$= \frac{\pi}{8}$