Math 526 Notes

Kellen Kanarios

November 15, 2023

Abstract

These are the notes for Math 395 taught by David Barret.

Contents

Chapter 1

TBD

1.1 Introduction

Lecture 2: Introduction

Read 1, 2, 3, A, B (note Theorem A.1) 30 Aug. 02:30

Let V, W be vector spaces over F. Assume $1+1\neq 0$. We will study functions $T: V \to W$. We define the Graph of T as follows

$$\begin{aligned}\text{Graph } T &= \{(\vec{v},\vec{w}) \in V\times W \mid \vec{w} = T\vec{v}\}\\ &= \{(\vec{v},T\vec{v}) \in V\times W \mid \vec{v}\in V\}\end{aligned}$$

Definition 1.1.1. T is affine if and only if Graph T is affine.

Special case: $T(\vec{0}) = \vec{0}$. Equivalently, $(\vec{0},\vec{0}) \in \text{Graph}T$. Then T is affine if and only if Graph T is a vector subspace. Furthermore, $\vec{v}_1, \vec{v}_2 \in V$, $t\in F$ implies

$$(\vec{v}_1+\vec{v}_2, T(\vec{v}_1)+T(\vec{v}_2)) = (\vec{v}_1, T(\vec{v}_1)) + (\vec{v}_2, T(\vec{v}_2)) \in \text{Graph}T$$

General case: $T: v\to w$ affine $\Leftrightarrow$ Graph T is affine $\Leftrightarrow$ Graph $\underbrace{T-(\vec{0},T(\vec{0}))}_{\{\vec{v},T(\vec{v}-T(\vec{0})\mid\vec{v}\in V\}=L}$ is a vector subspace.

Define

$$L(\vec{v}) \coloneqq T(\vec{v}) - T(\vec{0})$$

is linear function. T is of the form $T(v) = \tilde{T}(\vec{v}) + \vec{b}$ with $\tilde{T}$ linear.

Problem 1.1.1. Show $\tilde{T}$ is uniquely defined.

Often useful to extend $\mathbb{R}$ by formally adding the points $+\infty, -\infty$. Key rule: $\forall x\in\mathbb{R}$[1], $-\infty \le x \le +\infty$. Set of extended reals is $\mathbb{R}\cup\{-\infty,+\infty\} = [-\infty,+\infty]$. Now if $S\subset\mathbb{R}$ then the set of upper bounds for S has the form $[\underbrace{\beta}_{\sup S}, +\infty]$.

Example. For $S=\mathbb{Z}$, we have that $\sup S = +\infty$.

Example. For $S=\emptyset$, we have that $S = [-\infty,+\infty]$. Therefore, $\sup S = -\infty$.

Remark. $\sup S = -\infty \Leftrightarrow S = \emptyset$

[1] For x in extended reals too.

Note. $x \in S \Rightarrow \inf S \le x \le \sup S$. Therefore, $\inf S \le \sup S$. However, when $S = \emptyset$ we have that $\sup S = -\infty < +\infty = \inf S$.

$[-\infty, +\infty]$ not a group. We define $(+\infty) + 2 = +\infty$. Similarly, $(-\infty) + 17 = -\infty$. We refuse to define $(+\infty) + (-\infty)$. We also refuse to define $(\pm\infty) \cdot 0$.

Notation. Suppose $\alpha \in A$ we are given $S_\alpha \subset X$ i.e. we have a map $A \to \mathcal{P}(X)$.

$$\bigcup_{\alpha \in A} S_\alpha = \{x \in X \mid x \in S_\alpha \text{ for at least one } \alpha \in A\}$$

$$\bigcap_{\alpha \in A} S_\alpha = \{x \in X \mid x \in S_\alpha \text{ for all } \alpha \in A\}$$

Remark. For $A = \emptyset$, we have that

$$\bigcap_{\alpha \in A} S_\alpha = X.$$

Must specify a universal set X in advance.

1.2 Topological + Metric Spaces

Definition 1.2.1. A metric on a set X is a function $X \times X \to \mathbb{R}$ such that

1. $d(x, y) = d(y, x)$.
2. $d(x, y) \ge 0$, $d(x, y) = 0 \Leftrightarrow x = y$.
3. $d(x, z) \le d(x, y) + d(y, z)$.

Definition. A metric space is a set equipped with a metric.

Definition 1.2.2. Let X be a metric space. If $Y \subset X$ then $d|_{Y \times Y}$ is the induced metric.

Example. Most important examples are

1. $X = \mathbb{R}^n$, $d_{\text{eucl}}(\vec{x}, \vec{y}) = \sqrt{\sum_{j=1}^n (x_j - y_j)}$
2. $Y \subset \mathbb{R}^n$ with induced metric.

Definition 1.2.3. For $x_0 \in X$, $\epsilon > 0$

$$\mathcal{U}(x_0; \epsilon) = \{x \in X \mid d(x_0, x) < \epsilon\}$$

This is the ϵ-ball centred at x_0.

Notation. We will and you can also write $\mathcal{U}(x_0, \epsilon)$.

Definition. Consider $A \subset X^{\text{metric}}$.

Definition 1.2.4. x_0 is interior to A if and only if $\exists \epsilon > 0$ such that $\mathcal{U}(x_0, \epsilon) \subset A$.

Definition 1.2.5. x_0 is exterior to A if and only if $\exists \epsilon > 0$ such that $\mathcal{U}(x_0, \epsilon) \cap A = \emptyset$.

Definition 1.2.6. x_0 is a boundary point for A if and only if x_0 is neither interior nor exterior.

Notation. Int$A$ is the set of interior points. Ext$A$ is the set of exterior points. BdA is the set of boundary points.

Remark. We get that X is the disjoint union of Int$X$, Ext$X$, and BdX.

Definition 1.2.7. A is open if and only if $A = \text{Int}A$.

Proposition 1.2.1. This defines a topology on X.

Proof. See Section 3 in Munkres. ■

Remark. We will make use of the following additional facts

1. Each $\mathcal{U}(x_0, \epsilon)$ is open.
2. Int$A$ is the largest open subset of A.
3. A is closed if A^c is open.
4. $\overline{A}$ is smallest closed superset of A.
5. $\text{Bd}(X \setminus A) = \text{Bd}A$.
6. BdA is closed.

Definition 1.2.8. We define the closure of A as $\overline{A} = \text{Int}A \cup \text{Bd}A$.

Given $\{x_n\}$ sequence in X. Suppose that (X, d) is a metric space. Then $x_n \to x \Leftrightarrow \forall \epsilon > 0$, $\exists N$ such that $d(x_n, x) < \epsilon$ when $n > N$. x_n converges $\Leftrightarrow \exists x$ such that $x_n \to x$.

Problem 1.2.1. Suppose $x_n \to x$ and $x_n \to y$ then $x = y$.

Definition 1.2.9. Suppose $f : (X, d_X) \to (Y, d_Y)$. Then f is sequentially continuous if and only $x_n \to x$ implies $f(x_n) \to f(x)$.

Theorem 1.2.1. f is sequentially continuous if and only if f is continuous.

Proof. Last year did $X = \mathbb{R}$. ■

Lecture 5: TBD

30 Aug. 02:30

Proposition 1.2.2. For X top space we hace

$$f = (f_1, \dots, f_n) : X \to \mathbb{R}^n \text{ cont.} \Leftrightarrow \text{ each } f_j \text{ cont}$$

Example. $\mathbb{R}^2 \to \mathbb{R}$, $(x_1, x_2) \mapsto \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2} & (x_1, x_2) \neq (0,0) \\ 0 & (x_1, x_2) = (0,0) \end{cases}$ This is continuous function of x_1 for x_2

> fixed and vice versa. Not continuous $(1/n, 1/n) \to 0$ but $f(1/n, 1/n) \to 1/2$.

1.3 Directional Derivative

We have the following intuitions that may be a paradox?

- f cont. at a if and only if graph f is "almost horizontal" when magnified.
- f diff at a if and only if graph f is "almost affine" when magnified.

We are good if we use different magnifications for each. Namely, we only magnify in the horizontal direction for continuity and in both directions for differentiability.

We will now extend this to a vector space V. A dilation centered at $\vec{0}$ is just as follows

$$\begin{aligned} f : V &\to V \\ \vec{x} &\mapsto \lambda\vec{x} \end{aligned}$$

It then follows that a dilation centered at $\vec{p}$ is

$$\begin{aligned} f : V &\to V \to V \to V \\ \vec{x} &\mapsto \vec{x} - \vec{p} \mapsto \lambda(\vec{x} - \vec{p}) \mapsto \lambda(\vec{x} - \vec{p}) + \vec{p} = \lambda\vec{x} + (1 - \lambda)\vec{p} \end{aligned}$$

Given $f : V \to V$, $\text{Graph}(f) = \{(\vec{x}, f(\vec{x}) \in V \times W \mid \vec{x} \in V\}$. Dilate about $(\vec{a}, f(\vec{a})) = \vec{p}$ i.e.

$$(\vec{x}, f(\vec{x})) \mapsto (\lambda(\vec{x} - \vec{a}) + \vec{a}, \lambda(f(\vec{x}) - f(\vec{a})) + f(\vec{a}))$$

Set $t = 1/\lambda$, so that $\vec{x} = \vec{a} + t\vec{u}$. Then

$$\vec{u} = \frac{\vec{x} - \vec{a}}{t} = \lambda(\vec{x} - \vec{a})$$

Our Dilated graph becomes

$$\left(\vec{a} + \vec{u}, f(\vec{a}) + \frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t}\right)$$

What happens when $t \to 0$ (assuming field has topology)? We want this to be an affine set i.e. the graph of an affine function. Would need to have the form $T(\vec{u}) + b$. Set $\vec{u} = 0$, need $b = f(\vec{a})$.

Question reduces to does $\lim_{t\to 0} \frac{f(\vec{a}+t\vec{u})-f(\vec{a})}{t} = T(\vec{u})$?

> **Definition 1.3.1.** We call $f'(\vec{a}; \vec{u}) \coloneqq \lim_{t\to 0} \frac{f(\vec{a}+t\vec{u})-f(\vec{a})}{t}$ the **directional derivative** of f at $\vec{a}$ in direction $\vec{u}$.

1.3.1 Why is this not the basis of multivariate calculus?

First Reason

All $f'(\vec{a}; \vec{u})$ exist but not linear in $\vec{u}$.

Second Reason

Even though f' linear in u, we see that directional differentiability is not enough to ensure continuity.

> **Example.** Consider
>
> $$f(x, y) = \begin{cases} \frac{x^3}{y}, & y \neq 0 \\ 0, & y = 0 \end{cases}.$$
>
> Check: $f'(\vec{0}; \vec{u}) = \vec{0}$. However, $f(1/n, 1/n^4) = n \not\to 0$, so f not continuous at origin.

Third Reason

Chain Rule will fail without stronger assumptions.

1.3.2 Vector-valued function of a scalar

Definition 1.3.2. Consider $f : I^{\text{open interval in } \mathbb{R}} \to W^{\text{vector space}}$. Then we define $f'(x) := \lim_{t\to 0} \frac{f(x+t)-f(x)}{t}$.

Remark. Choose W to be a normed vector space. If $\dim W < \infty$ all norms on W give same topology on W. In fact, $\mathrm{Id} : (W, d_1) \to (W, d_2)$ is bi-Lipschitz.

Proof. HW 3 or 4. ∎

1.3.3 General Case

For $\vec{a}, \vec{u} \in V$, define $g_{\vec{a},\vec{u}} : \mathbb{R} \to V$ by $t \mapsto \vec{a} + t\vec{u}$.

Exercise. $g'_{\vec{a},\vec{u}}(t) = \vec{u},\ \forall t \in \mathbb{R}$.

We now return to $f : V \to W$. We have

$$(f \circ g_{\vec{a},\vec{u}})'(0) = \lim_{t\to\infty} \frac{(f \circ g_{\vec{a},\vec{u}})(t) - (f \circ g_{\vec{a},\vec{u}})(0)}{t} \tag{1.1}$$

Exercise. Show (1.1) is just $f'(\vec{a}; \vec{u})$.

Remark. Directional derivative reduces to derivative of vector-valued function of scalar.

Remark. Best to assume $\vec{a} \in \mathrm{Int}(\mathrm{dom}(f))$.

Back to dilated graphs

Dilated Graph f about $(\vec{a}, f(\vec{a}))$, take limit, dilated graphs converge pointwise to

$$\begin{aligned}\{(\vec{a} + \vec{u}, f(\vec{a}) + f'(\vec{a}; \vec{u})) \mid \vec{u} \in V\} &= \{(\vec{y}, f(\vec{a}) + f'(\vec{a}; \vec{y} - \vec{a})) \mid \vec{y} \in V\} \\ &= \mathrm{Graph}(\vec{y} \mapsto f(\vec{a}) + f'(\vec{a}; \vec{y} - \vec{a}))\end{aligned}$$

To get good theory we need the following things

1. $f'(\vec{a}; \vec{u})$ linear in $\vec{u}$.
2. Also want T continuous.

Facts:

(a) For $T : V \to W$ linear, TFAE
   - (i) T is continuous,
   - (ii) $\exists M \in [0, +\infty)$ such that $\|T\vec{v}\| \leq M\|\vec{v}\|, \forall \vec{v} \in V$

(b) Both automatically hold when $\dim V < \infty$.

Explicitly, we want $f(\vec{a}) + T(\vec{y} - \vec{a}) \approx f(\vec{y})$.

Lecture 6: TBD

13 Sep. 02:30

Definition 1.3.3. $\operatorname{Hom}(V, W) = $ (set of linear $T : V \to W$)

Remark. For $T \in \operatorname{Hom}(V, W)$, TFAE

(a) T is continous at $\vec{0}$

(b) T is continous on V.

(c) T is Lipschitz.

(c) $\exists M > 0$ such that $\|T(\vec{v})\| \leq M\|\vec{v}\|$ for all $\vec{v} \in V$.

Definition 1.3.4. $B(V, W) := \{T \in \operatorname{Hom}(V, W) : T \text{ satisfies one of above}\}$

Remark. If $\dim V < \infty$ then $\operatorname{Hom}(V, W) = B(V, W)$.

From last lecture, for some $f : A^{\text{open in } V} \to W$, we are interested in approximating f near a by affine map. Namely,

$$f(\vec{a} + \vec{h}) \approx f(\vec{a}) + T(\vec{h})$$

Exercise. All affine $F(\vec{a} + \vec{u})$ such that $F(\vec{a}) = f(\vec{a})$ have this form.

Concretely, want

$$\frac{f(\vec{a} + \vec{h}) - f(\vec{a}) - T(\vec{h})}{\left\|\vec{h}\right\|} \to 0 \text{ as } \vec{h} \to \vec{0}$$

Definition 1.3.5. We say $Df(\vec{a}) = T$ if for $T \in \operatorname{Hom}(V, W)$

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

Definition 1.3.6. f is differentiable at $\vec{a}$ if and only if $\exists T \in B(V, W)$ such that $Df(\vec{a}) = T$.

Problem 1.3.1. How do we find a candidate for T?

Answer. From Monday, if T exists then

$$T(\vec{h}) = f'(\vec{a}; \vec{h})$$

⊛

Proposition 1.3.1. If f differentiable at $\vec{a}$ then f is continuous at $\vec{a}$.

Remark. This fails if we allow discontinuous T.

Proof. It suffices to show that $f(\vec{x}) \to f(\vec{a})$ as $\vec{x} \to \vec{a}$. Write $\vec{x} = \vec{a} + \vec{h}$, $\vec{h} = \vec{x} - \vec{a}$. Now we have the equivalent statement that we need $f(\vec{a}+\vec{h})-f(\vec{a}) \to \vec{0}$ as $\vec{h} \to 0$. Note that

$$+\ \mathrm{T}()\ f(\vec{a}+\vec{h})-f(\vec{a}) = \left\|\vec{h}\right\| \underbrace{\frac{f(\vec{a}+\vec{h})-f(\vec{a})-T(\vec{h})}{\left\|\vec{h}\right\|}}\star$$

By triangle inequality,

$$\left\|f(\vec{a}+\vec{h})-f(\vec{a})\right\| \leq \left\|\vec{h}\right\| \|\star\| + \left\|T(\vec{h})\right\|$$

All terms in the RHS go to $\vec{0}$ as $\vec{h} \to \vec{0}$. Thus, we are done. ■

Example (Special Case A). $W = \mathbb{R}^n$. Then $f = (f_1, \ldots, f_n)$ or better... $f = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$

Proposition 1.3.2. f differentiable at $\vec{a}$ if and only if each f_i differentiable at $\vec{a}$. Moreover,

$$Df(\vec{a})(\vec{u}) = \begin{bmatrix} Df_1(\vec{a})(\vec{u}) \\ \vdots \\ Df_n(\vec{a})(\vec{u}) \end{bmatrix}$$

Example (Special Case B). $V = \mathbb{R}^m$. Then

$$Df(\vec{a}) \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} = Df(\vec{a})(\sum u_j \vec{e}_j) = \sum u_j Df(\vec{a})(\vec{e}_j) = \sum u_j f'(\vec{a}; \vec{e}_j)$$

We denote $D_j f(\vec{a}) := f'(\vec{a}; \vec{e}_j)$. Then

$$Df(\vec{a})(\vec{u}) = \begin{bmatrix} D_1 f(\vec{a}) & \cdots & D_m f(\vec{a}) \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$$

Example (Special Case C). $V = \mathbb{R}^m$, $W = \mathbb{R}^n$. Then each $D_j f(\vec{a})$ is a column vector in $\mathbb{R}^n$. Therefore,

$$\begin{bmatrix} D_1 f(\vec{a}) & \cdots & D_m f(\vec{a}) \end{bmatrix}$$

is an $n \times m$ matrix representing $Df(\vec{a})$, where the (j,k)-entry is $D_k f_j(\vec{a})$.

Case 1: $m = 1$.

We discussed this on Monday.

Case 2: $n = 1$.

Physical prototype: temperature function/altitude function.

$$Df(\vec{a}) = \begin{bmatrix} D_1 f(\vec{a}) & \cdots & D_m f(\vec{a}) \end{bmatrix} \in \mathbb{R}^m_{\text{row}}$$

This is the gradient of f denoted $\operatorname{grad} f(\vec{a})$.

1.4 Chain Rule

Suppose $f : A^{\text{open subset of } V} \to W$ with f differentiable at $\vec{a}$. Additionally, suppose $f(\vec{a}) = \vec{b} \in B^{\text{open subset of } W}$. Now suppose we have $g : B^{\text{open subset of } W} \to Z$, where Z is a normed vector space and g is differentiable at $\vec{b}$. Then

$$\begin{aligned}(g \circ f)(\vec{x}) &\approx g(\vec{b}) + Dg(\vec{b})(f(\vec{x}) - f(\vec{a})) \\ &\approx (g \circ f)(\vec{a}) + (Dg(\vec{b}) \circ Df(\vec{a}))(\vec{x} - \vec{a})\end{aligned}$$

This suggests that

$$D(g \circ f)(\vec{a}) = Dg(\vec{b}) \circ Df(\vec{a})$$

This is the multivariable chain rule.

Theorem 1.4.1. Given f, g as above the following holds

$$D(g \circ f)(\vec{a}) = Dg(\vec{b}) \circ Df(\vec{a})$$

if f differentiable at $\vec{a}$ and g differentiable at $\vec{b}$.

Example. If we return to the single-variable calculus world

$$\underbrace{\mathbb{R}}_{\ni a} \xrightarrow{f} \underbrace{\mathbb{R}}_{\ni b = f(a)} \xrightarrow{g} \mathbb{R}$$

From single-variable calculus,

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a) = g'(b) \cdot f'(a)$$

Without a for some reason,

$$(g \circ f)' = (g' \circ f) \cdot f'$$

Using product rule,

$$(g \circ f)'' = (g' \circ f)' \cdot f' + (g' \circ f) \cdot f''$$

Lecture 7: TBD

15 Sep. 02:30

Proof of Theorem 1.4.1. We want to show that

$$\lim_{\vec{h} \to \vec{0}} \frac{g(f(\vec{a} + \vec{h})) - g(f(\vec{a})) - Dg(\vec{b})(Df(\vec{a})(\vec{h}))}{\|\vec{h}\|} = \vec{0}. \tag{1.2}$$

Define

$$F(\vec{h}) := = \begin{cases} \frac{f(\vec{a}+\vec{h}) - f(\vec{a}) - Df(\vec{a})(\vec{h})}{\|\vec{h}\|}, & \text{if } \vec{h} \neq 0 \\ \vec{0}, & \text{if } \vec{h} = 0 \end{cases}$$

Then F continuous at $\vec{0}$ and $\lim_{\vec{h} \to \vec{0}} F(\vec{h}) = \vec{0}$. Note that

$$f(\vec{a} + \vec{h}) - f(\vec{a}) = Df(\vec{a})(\vec{h}) + \|\vec{h}\| F(\vec{h})$$

Similarly,

$$g(\vec{a} + \vec{h}) - g(\vec{a}) = Dg(\vec{a})(\vec{h}) + \|\vec{h}\| G(\vec{h})$$

with $G(\vec{0}) = \vec{0}$, G continuous at $\vec{0}$. Introduce

$$\vec{k} = f(\vec{a} + \vec{h}) - f(\vec{a}) = Df(\vec{a})(\vec{h}) + \|\vec{h}\|F(\vec{h}),$$

so that $g(\vec{h} + \vec{k}) = g(f(\vec{a} + \vec{h}))$. Then

$$\begin{aligned}
(1.2) &= \frac{g(\vec{b} + \vec{k}) - g(\vec{b}) - Dg(\vec{b})(Df(\vec{a})(\vec{h})}{\|\vec{h}\|} \\
&= \frac{Dg(\vec{b})(\vec{k}) + \|\vec{k}\|G(\vec{k}) - Dg(\vec{b})(Df(\vec{a})(\vec{h}))}{\|\vec{h}\|} \\
&= \frac{Dg(\vec{b})Df(\vec{a})(\vec{h}) + \|\vec{h}\|Dg(\vec{b})F(\vec{h}) + \|\vec{k}\|G(\vec{k}) - Dg(\vec{b})Df(\vec{a})(\vec{h})}{\|\vec{h}\|} \\
&= Dg(\vec{b})F(\vec{h}) + \frac{\|\vec{k}\|}{\|\vec{h}\|}G(\vec{k})
\end{aligned}$$

From above, we know that $F(\vec{h}) \to 0$. Since $Dg(\vec{b})$ is continuous, it follows that $Dg(\vec{b})F(\vec{h}) \to 0$. Similarly, we have that $\vec{k} \to 0$ and $\vec{h} \to 0$, so that $G(\vec{k}) \to 0$. To show that the product goes to 0, we must show that $\frac{\|\vec{k}\|}{\|\vec{h}\|}$ is bounded. Note that

$$\frac{\|\vec{k}\|}{\|\vec{h}\|} = \|Df(\vec{a})\frac{\vec{h}}{\|\vec{h}\|} + F(\vec{h})\|$$

Note $F(\vec{h})$ is continuous at $\vec{h} = 0$. Therefore, it is bounded in a neighborhood of $\vec{0}$. Since $Df(\vec{a})$ is bounded,

$$\|Df(\vec{a})\frac{\vec{h}}{\|\vec{h}\|}\| \leq M\|\frac{\vec{h}}{\|\vec{h}\|}\| = M$$

■

Lecture 8: TBD

18 Sep. 02:30

Exercise. $D(f_1, f_2)(\vec{a}) = (Df_1(\vec{a}), Df_2(\vec{a})) \in B(V, W \times W)$

Example. $W \times W \xrightarrow{\alpha} W$ defined by $(\vec{w}_1, \vec{w}_2) \mapsto \vec{w}_1 + \vec{w}_2$. Then $D_\alpha(\vec{w}_1, \vec{w}_2) = \alpha$.

Example. Consider $\underbrace{f_1 + f_2}_{=\alpha\circ(f_1, f_2)} : A \to W$. By Chain rule, $f_1 + f_2$ differentiable at $\vec{a}$. Then

$$\begin{aligned}
D(f_1 + f_2)(\vec{a}) &= D_\alpha(f_1(\vec{a}), f_2(\vec{a})) \circ (D_1 f_1(\vec{a}), D_2 f_2(\vec{a})) \\
&= Df_1(\vec{a})(\vec{h}) + Df_2(\vec{a})(\vec{h})
\end{aligned}$$

Example. $\mathbb{R}^2 \xrightarrow{\mu} \mathbb{R}$ defined by $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1 x_2$. If it exists

$$D_\mu \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{bmatrix} D_1\mu \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & D_2\mu \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{bmatrix} = \begin{bmatrix} x_2 & x_1 \end{bmatrix}$$

Example. Suppose $A \overset{f_1}{\underset{f_2}{\rightrightarrows}} \mathbb{R}$ differentiable at $\vec{a} \in \mathbb{R}$. Consider $\mu \circ \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$ defined by $\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \mapsto f_1 f_2$. Then by chain rule $f_1 f_2$ differentiable.

$$D(f_1, f_2)(\vec{a}) = D_\mu \begin{bmatrix} f_1(\vec{a}) \\ f_2(\vec{a}) \end{bmatrix} \circ \begin{bmatrix} Df_1(\vec{a}) \\ Df_2(\vec{a}) \end{bmatrix} = f_2(\vec{a})Df_1(\vec{a}) + f_1(\vec{a})Df_2(\vec{a})$$

Example. Extend to $f_1 : A \to \mathrm{Mat}(m, n, \mathbb{R})$ and $f_2 : A \to \mathrm{Mat}(n, p, \mathbb{R})$. Then $f_1 f_2 : A \to \mathrm{Mat}(n, p, \mathbb{R})$.

$$D(f_1 f_2)(\vec{a})(\vec{h}) = Df_1(\vec{a})(\vec{h}) \cdot f(\vec{a}) + f_1(\vec{a}) \cdot Df_2(\vec{a})(\vec{h})$$

Example. f_1, f_2 $\mathbb{C}$-valued identify $a + ib$ with $\begin{bmatrix} a \\ b \end{bmatrix}$ with $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

Given $\vec{a} \in A^{\text{osso } V}$, $\vec{h} \in V$, $Df(\vec{a})(\vec{h}) \in W$. We've been insisting that this be continuous with respect to $\vec{h}$. What about dependence on $\vec{a}$? If f differentiable at each $\vec{a} \in A$, we get a map from $Df : A \to B(V, W)$ defined by $\vec{a} \mapsto Df(\vec{a})$.

Definition 1.4.1. f is continuously differentiable on A if and only if Df exists and is continuous on A. We will use the notation $f \in C^1$.

Remark. Continuity of Df is not automatic. Even when $V = W = \mathbb{R}$. See Section 6 #2.

Example (Special Case). $V = \mathbb{R}^m$, $W = \mathbb{R}^n$. $B(V, W) = \mathrm{Mat}(n, m, \mathbb{R})$. Then $f \in C^1$ if and only if f differentiable at each $\vec{a} \in A$ and $Df : A \to \mathrm{Mat}(n, m, \mathbb{R})$ is continuous. Equivalently, we want f is differentiable at each $\vec{a} \in A$ and each $D_k f_j = f_j'(\vec{a}, \vec{e}_k)$ is continuous on A.

Theorem 1.4.2. Given $f : A^{\text{osso } \mathbb{R}^m} \mathbb{R}^n$ all $D_k f_j$ exist and are continuous on A then $f \in C^1(A, \mathbb{R}^n)$. In particular, f differentiable at each $\vec{a} \in A$.

Proof. Suffices to show f is differentiable at each $\vec{a} \in A$. From Wednesday, it suffices to consider $n = 1$. For clarity, we will just consider $m = 2$ (see textbook for higher dimensions). Fix $\vec{a} \in A \subset \mathbb{R}^2$, consider $\vec{h}$ small.

Remark. Want to see if $f(\vec{a} + \vec{h}) - f(\vec{a})$ is well approximated by linear function of $\vec{h}$.

By MVT, $\exists \vec{p}$ so that $f(a_1 + h_1, a_2) - f(a_1, a_2) = D_1 f(\vec{p}) h_1$. Similarly, $\exists \vec{q}$, so that $f(a_1, a_2 + h_2) - f(a_1, a_2) = D_2 f(\vec{q}) h_2$. Then $f(\vec{a} + \vec{h}) - f(\vec{a}) = D_1 f(\vec{p}) h_1 + D_2 f(\vec{a}) h_2$. Recall that if the derivative exists it must be the gradient. Therefore, our goal is now to show that

$$\frac{f(\vec{a} + \vec{h}) - f(\vec{a}) - (D_1 f(\vec{a}) h_1, D_2 f(\vec{a}) h_2)}{\|\vec{h}\|} \to 0 \text{ as } \vec{h} \to 0 \qquad (\star)$$

However,

$$(\star) = (D_1 f(\vec{p}) - D_1 f(\vec{a})) \frac{h_1}{\|\vec{h}\|} + (D_2 f(\vec{q}) - D_2 f(\vec{a})) \frac{h_2}{\|\vec{h}\|} \to 0$$

Note that $\frac{h_I}{\|\vec{h}\|} \leq 1$. Therefore, we only need to show the differences go to 0. This follows from our assumptions on the continuity of D_i. Since $D_1 f$ and $D_2 f$ are continuous at $\vec{a}$, as $h \to 0$, we have that $\vec{p} \to \vec{a}$ and $D_1 f(\vec{p}) - D_1 f(\vec{a})) \to 0$. ■

Next Lecture Preview: Given $f \in C^1(A^{\text{osso } V}, W)$ leads to $Df : A \to B(V, W)$ continuous. This might be differentiable. If so, $D^2 f := D(Df) : A \to B(V, B(V, W))$. $D^2 f$ might or might not be continuous. If so, we say that $f \in C^2$.

Lecture 9: TBD

Give $f : A^{\text{osso } V} \to W$, we have some good things that can happen. In increasing strength, we have: 20 Sep. 02:30

(a) All $Df(\vec{a}; \vec{h})$ exist for $\vec{a} \in A, \vec{h} \in V$.

(b) f differentiable at all $\vec{a} \in A$.

(c) $Df : A \to B(V, W)$ is continuous.

- From Monday, this happens if and only if all $D_k f_j$ exist and are continuous.

(d) $Df : A \to B(V, W)$ is differentiable.

- Get $D^2 f = D(Df) : A \to B(V, B(V, W))$.

(e) $D^2 f$ is continuous.

Definition 1.4.2. $f \in C^2 \Leftrightarrow Df$ differentiable at each $\vec{a} \in A$ and $D^2 f$ is continuous. $C^r(A, W)$ is similar for $r \in \mathbb{N}$. We will denote $C^r(A) = C^r(A; \mathbb{R})$.

Theorem 1.4.3. Given $f \in C^2(A^{\text{osso } \mathbb{R}^2}, \mathbb{R})$. Then

$$D_2 D_1 f = \lim_{(h,k) \to (0,0)} \frac{f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, h)}{hk}.$$

Proof. Let $\phi(s) = f(s, b+k) - f(s, b)$ defined on open interval $\supset [a, a+h]$. Then

$$\phi'(s) = D_1 f(s, b+k) - D_1 f(s, b)$$

This can be seen by considering the following mapping

$$s \mapsto (s, b+k) \mapsto f(s, b+k)$$

The first mapping has derivative $\begin{bmatrix} D_1 f & D_2 f \end{bmatrix}$. The second mapping has derivative $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. So by the chain rule, the composition has derivative $D_1 f$. Then the numerator can be written as

$$\begin{aligned}
\phi(a+h) - \phi(a) &= \phi'(s_0) \cdot h \\
&= (D_1 f(s_0, b+k) - D_1 f(s_0, b)) \cdot h \\
&= D_2 D_1 f(s_0, t_0) \cdot k \cdot h \\
&= D_2 D_1 f(s_0, t_0) \to D_2 D_1 f(a, b) \text{ as } (h, k) \to (0, 0)
\end{aligned}$$

■

Corollary 1.4.1. Given $f \in C^2(A^{\text{osso } \mathbb{R}^2}, \mathbb{R})$. Then $D_2 D_1 f = D_1 D_2 f$.

Notation. For $f : A^{\text{osso } V} \to W$ differentiable, we write $D_{\vec{u}} f : A \to W$ for the mapping $\vec{a} \mapsto f'(\vec{a}, \vec{u}) = Df(\vec{a})(\vec{u})$.

Corollary 1.4.2. $f \in C^2(A^{\text{osso } V}, \mathbb{R}) \Rightarrow D_{\vec{u}_1} D_{\vec{u}_2} f = D_{\vec{u}_2} D_{\vec{u}_1} f$.

Proof. Apply chain rule twice to $(x_1, x_2) \mapsto f(\vec{a} + x_1 \vec{u}_1 + x_2 \vec{u}_2)$ i.e. make a sequence of maps $\mathbb{R}^2 \to V \to \mathbb{R}$. ■

Study. We can do (I think?)

$$D_{\vec{u}_1} D_{\vec{u}_2} f = \lim_{(h,k)\to(0,0)} \frac{f(\vec{a}+h\vec{u}_1+k\vec{u}_2) - f(\vec{a}+h\vec{u}_1) - f(\vec{a}+k\vec{u}_2) + f(\vec{a})}{hk}$$

■

Corollary 1.4.3. $f \in C^2(A^{\text{osso } V}, \mathbb{R}^n) \Rightarrow D_{\vec{u}_1} D_{\vec{u}_2} f = D_{\vec{u}_2} D_{\vec{u}_1} f$.

Proof. Apply above to each f_j and reassemble. ■

Remark. Also works for $f \in C^2(A, W)$

1. If $\dim W < \infty$: use equivalence of norms that will be proven on HW4.
2. If $\dim W = \infty$: requires additional tool from HW4 study exercise.

Corollary 1.4.4. Suppose $f \in C^r$, consider $D_{\vec{u}_1} D_{\vec{u}_2} \dots D_{\vec{u}_r} f$. Then

- Can interchange $\vec{u}_j$ with $\vec{u}_{j+1}$.
- Can arbitrarily permute the $\vec{u}_j$.

Remark. Existence of $D^r f$ is not enough for the above corollary.

Exercise. For $f \in C^r(\mathbb{R}^m, \mathbb{R})$ how many distinct $D_{j_1} \dots D_{j_r} f$ are there?

Answer. Hint: rewrite as $D_1^{\alpha_1} \cdots D_m^{\alpha_m} f$ where $\alpha_j \in \mathbb{N} \cup \{0\}$ and $\sum_{j=1}^m \alpha_j \le r$. Preview: notate this as $D^{(\alpha_1, \dots, \alpha_m)} f$ called a "multi-index". ⊛

1.5 Inverse Function Theorem

1.5.1 Inverse Functions

Given $A^{\text{open}} \subset V$, $B^{\text{open}} \subset W$. Suppose $f : A \to B$ differentiable at $\vec{a}$, $g : B \to A$ differentiable at $\vec{b} = f(\vec{a})$ and $g \circ f = \mathrm{Id}_A$. Then

$$Dg(\vec{b}) \circ Df(\vec{a}) = \mathrm{Id}\,.$$

Therefore, $Dg(\vec{b})$ is a left inverse for $Df(\vec{a})$. It follows that $\dim V \le \dim W \in \mathbb{N} \cup \{0, +\infty\}$.

(1) If also $f \circ g = \mathrm{Id}_B$ then $Dg(\vec{b})$ is a two-sided inverse for $Df(\vec{a})$ and $\dim V = \dim W$.

(2) If we assume $\dim V = \dim W < \infty$ then $Dg(\vec{b})$ is a two-sided inverse for $Df(\vec{a})$ and $f \circ g = \mathrm{Id}_B$.

(3) If A, B as above and $\dim V, \dim W < \infty$ with $f : A \to B$ cont, $G : B \to A$ cont, and $g \circ f = \mathrm{Id}_A$ and $f \circ g = \mathrm{Id}_B$ then $\dim V = \dim W$.

(4) $\exists \mathbb{R} \to \mathbb{R}^2$ continuous surjection. "Space filling curve".

Lecture 10: TBD

Remark. If you replace continuous in (3) with differentiable then we did this on Wednesday. 22 Sep. 02:30

Definition 1.5.1. A *homeomorphism* is a continuous bijection $f : A \to B$ such that f^{-1} is continuous.

Example. $f : [0, 2\pi) \to S^1 := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ given by $f(t) = (\cos t, \sin t)$ is a continuous bijection but not a homeomorphism. f^{-1} is not continuous because $f^{-1}(\cos 1/n, -\sin 1/n) = 2\pi - 1/n$ but $f^{-1}(1, 0) = 0$. Therefore, $f^{-1}(\cos 1/n, -\sin 1/n) \not\to f^{-1}(1, 0)$ and f^{-1} is not continuous.

Definition 1.5.2. A C^r-*diffeomorphism* is a C^r bijection $f : A^{\text{osso } V} \to B^{\text{osso } W}$ such that f^{-1} is C^r.

Example. $f : \mathbb{R} \to \mathbb{R}$ defined by $t \mapsto t^3$ is C^∞ and a bijection. $f^{-1} : \mathbb{R} \to \mathbb{R}$ defined by $t \mapsto \sqrt[3]{t}$ is not differentiable. Thus, f is not a C^r diffeomorphism.

Note. Composition of C^r diffeomorphisms is a C^r diffeomorphism.

1.5.2 Inverse Function Theorem

Theorem 1.5.1. Given $\vec{a} \in A^{\text{open}} \subset \mathbb{R}^n$, $f \in C^r(A, \mathbb{R}^n)$, $Df(\vec{a})$ is invertible. Then there exists open $U \subset A$ containing $\vec{a}$ such that $f|_U$ is a C^r diffeomorphism i.e. f maps U bijectively to open set and f^{-1} is C^r.

Remark. Also works in dimension case if V, W complete i.e. Banach spaces.

Definition 1.5.3. A *Banach* space is a complete normed vector space.

Example. $A = \left\{ \begin{pmatrix} r \\ \theta \end{pmatrix} \in \mathbb{R}^2 \mid r > 0 \right\}$, $f \begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$ is C^∞. Note

$$Df \begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}$$

Jacobian determinant $:= \det Df \begin{pmatrix} r \\ \theta \end{pmatrix} = r$. $Df \begin{pmatrix} r \\ \theta \end{pmatrix}$ is invertible for every $\begin{pmatrix} r \\ \theta \end{pmatrix} \in A$ but f not injective on A. $f(A) = \mathbb{R}^2 \setminus \{\vec{0}\}$. Get **local** C^∞ inverses, no global inverse.

Note. A few notes:

(1) $E := \{\vec{x} \in A \mid Df(\vec{x}) \text{ is invertible}\} = \{\vec{x} \in A \mid \det Df(\vec{x}) \neq 0\}$ is an open set containing $\vec{a}$.

(2) Theorem does not assume that $E = A$... but it could!

More about C^r functions:

Definition 1.5.4. $C^\infty(A, W) := \bigcap_{r \in \mathbb{N}} C^r(A, W)$.

For $f : A^{\text{osso } V} \to B^{\text{osso } W}$, $g : B \to Z$ differentiable, can write chain rule as

$$((Dg) \circ f) \circ Df$$

If we have stronger condition $f, g \in C^1$ then all maps on RHS are continuous i.e. $s \circ g \circ f \in C^1$. For $f \in C^2$, we have $D^2 f : A \to B(V, B(V, W))$.

Example. Consider $f : \vec{x} \mapsto T\vec{x} + \vec{b}$. Then $Df(\vec{x}) = T$ and $D^2 f(\vec{x}) = 0$.

In above settings,

$$f, g \in C^2 \Rightarrow Df, Dg \in C^1 \Rightarrow D(g \circ f) \in C^1 \Rightarrow g \circ f \in C^2$$

Exercise. Try the following exercises:

(a) Prove that $f, g \in C^r$ then $g \circ f \in C^r$.

(b) Prove that $f \in C^r(A, \mathbb{R} \setminus \{0\})$ then $\frac{1}{f} \in C^r$.

(c) Prove that $f, g \in C^r$ then $f + g, fg \in C^r$.

1.5.3 Preliminaries

Let $T_{\vec{a}} : \vec{x} \mapsto \vec{x} + \vec{a}$. Then $DT_{\vec{a}} = \mathrm{Id}$. Consider

$$\psi \coloneqq Df(\vec{a})^{-1} \circ T_{-f(\vec{a})} \circ f \circ T_{\vec{a}}$$

Check:

- $\psi(\vec{0}) = \vec{0}$.
- $D\psi(\vec{0}) = \mathrm{Id}$.
- $f = T_{f(\vec{a})} \circ Df(\vec{a}) \circ \psi \circ T_{-\vec{a}}$.

Suffices to show ψ is a diffeomorphism on some open set containing $\vec{0}$.

Lecture 11: TBD

25 Sep. 02:30

Proof. Let $h = f - \mathrm{Id}$, so that $Dh = Df - I$. Then $Dh(\vec{0}) = 0$ i.e. each $D_j h_k(\vec{0}) = 0$. Fix $\mu > 0$. There is $\delta > 0$ such that $|D_j h_k(\vec{x})| < \mu$ for $\vec{x} \in U(\vec{0}, \delta)$. If $\vec{x} \in U(\vec{0}, \delta)$ then

$$Dh(\vec{x})(\vec{u}) = \begin{pmatrix} \sum_j D_j h_1(\vec{x}) \cdot u_j \\ \vdots \\ \sum_j D_j h_n(\vec{x}) \cdot u_j \end{pmatrix}$$

Taking the norm, we have that

$$\|Dh(\vec{x})(\vec{u})\| \leq n^{3/2} \cdot \mu \cdot |\vec{u}| \leq n^{3/2} \cdot \mu \cdot \|\vec{u}\|$$

Therefore,

$$\|Dh\| \leq n^{3/n} \cdot \mu \text{ on } U(\vec{0}, \delta) \tag{1.3}$$

We will show that (1.3) implies that $\|h\|(\vec{x}) = \left\|h(\vec{x}) - h(\vec{0})\right\| \leq n^{3/2}\mu\|\vec{x}\|$. Fix ϵ with $0 < \epsilon < 1$. Let $\mu = \frac{\epsilon}{n^{3/2}}$. From above, we have that $\exists \delta > 0$ such that $\vec{x} \in U(\vec{0}, \delta)$ implies $Dh(\vec{0}) < n^{3/2}\mu = \epsilon$. From HW5 #1, $\|h(\vec{x}) - h(\vec{y})\| \leq \epsilon\|\vec{x} - \vec{u}\|$ for $\vec{x}, \vec{y} \in U(\vec{0}, \delta)$ convex.

Claim. f is bi-Lipschitz on $U(\vec{0}, \delta)$.

Proof. We have that

$$\|f(\vec{x}) - f(\vec{y})\| = \|\vec{x} - \vec{y} + h(\vec{x}) - h(\vec{y})\| \leq \|\vec{x} - \vec{y}\| + \|h(\vec{x}) - h(\vec{y})\| = (1 + \epsilon)\|\vec{x} - \vec{y}\|$$

We also have that

$$(1 - \epsilon)\|\vec{x} - \vec{y}\| \leq \|\vec{x} - \vec{y}\| - \|h(\vec{x}) - h(\vec{y})\| \leq \|f(\vec{x}) - f(\vec{y})\|$$

In particular, $f(U(0, \delta)) \subset U(\vec{0}, (1 + \epsilon)\delta)$ and f is injective on $U(\vec{0}, \delta)$. ■

Let $\psi_{\vec{y}} : \vec{x} \mapsto \vec{y} - h(\vec{x})$. Pick $0 < \tilde{\delta} < \delta$. Then $\vec{y} \in U(\vec{0}, (1 - \epsilon)\tilde{\delta})$, $\vec{x} \in U(0, \tilde{\delta})$. Therefore,

$$\|\psi_{\vec{y}}(\vec{x})\| \leq (1 - \epsilon)\tilde{\delta} + \epsilon\tilde{\delta} = \tilde{\delta}$$

It follows that $\psi_{\vec{y}} : U(\vec{0}, \tilde{\delta}) \to U(\vec{0}, \tilde{\delta})$, with

$$\|\psi_{\vec{y}}(\vec{x}) - \psi_{\vec{y}}(\vec{z})\| = \|h(\vec{x_1}) - h(\vec{x_2})\| \leq \epsilon\|\vec{x} - \vec{z}\|$$

Since $\epsilon < 1$, this is a contraction. To apply the contraction mapping theorem, we must have a mapping on a complete metric space. To do this, we can simply replace all U's by $\overline{U}$. Now we can apply contraction mapping theorem such that $\exists! \vec{x} \in \overline{U(\vec{0}, \tilde{\delta})}$, solving $\psi_{\vec{y}}(\vec{x}) = \vec{x}$. Note that this is equivalent to

$$y - f(\vec{x}) + \vec{x} = \vec{x}$$

i.e. $\vec{y} = f(\vec{x})$. Thus, f is locally surjective. ■

Remark. It is often useful to rewrite analysis problems as finding a fixed point.

Summary:

1. $\|h(\vec{x}) - h(\vec{y})\| \leq \epsilon\|\vec{x} - \vec{y}\|$ for $\vec{x}, \vec{y} \in U(\vec{0}, \delta)$.
2. $(1 - \epsilon)\|\vec{x} - \vec{y}\| \leq \|f(\vec{x}) - f(\vec{y})\| \leq (1 + \epsilon)\|\vec{x} - \vec{y}\|$.
3. f is injective on $U(\vec{0}, \delta)$.
4. $U(\vec{0}, (1 - \epsilon)\tilde{\delta}) \subset f\left(\overline{U(\vec{0}, \tilde{\delta})}\right) \subset f(A)$.
   - Take union over $\tilde{\delta} \in (0, \delta)$ to get $U(\vec{0}, (1 - \epsilon)\delta) \subset f(U(\vec{0}, \delta))$.

5. In particular, $\vec{0} \in \operatorname{Int}(f(A))$.

Theorem 1.5.2 (Cousin of Inverse Function Theorem)**.** Given $E^{\text{open}} \subset \mathbb{R}^n$, $f \in C^1(E, \mathbb{R}^n)$, $\det Df \neq 0$ on E then

(1) $\vec{a} \in E \Rightarrow f(\vec{a}) \in \operatorname{Int}(f(E))$.

(2) $f(E)$ is open in $\mathbb{R}^n$.

(3) $f : E \to f(E)$ is an open map.

Definition 1.5.5. f is an *open map* if whenever U is open $f(U)$ is open.

Proof of (1). Apply above to $\psi = Df(\vec{a})^{-1} \circ T_{-f(\vec{a})} \circ f \circ T_{\vec{a}}$. $\vec{0} \in \operatorname{Int}(\text{image of } \psi)$. Apply $Df(\vec{a})$, $\vec{0} \in \operatorname{Int}(\text{image of } T_{-f(\vec{a})} \circ f \circ T_{\vec{a}})$. Apply $T_{f(\vec{a})}$, $f(\vec{a}) \in \operatorname{Int} \underbrace{(\text{image of } f \circ T_{\vec{a}})}_{\text{image of } f}$ ■

Proof of (2). This is immediate. ■

Proof of (3). For $U^{\text{open}} \subset E$, apply (2) to $f|_U$. ■

Remark. So we have a local f^{-1}, we want that f^{-1} is C^r.

Proof of $f^{-1} \in C^r$. First we will do $r = 1$. Want to show $\rho := f^{-1}$ is C^1. Define $\vec{b} := f(\vec{a})$ and $M := Df(\vec{a})$. To show f^{-1} is differentiable, need

$$\frac{\rho(\vec{b} + \vec{k}) - \rho(\vec{b}) - T \cdot \vec{k}}{\|\vec{k}\|} \to 0 \tag{1.4}$$

Try $T = Df(\vec{u})^{-1}$. Note that

$$\vec{h} = \rho(\vec{b} + \vec{k}) - \rho(\vec{b})$$

Therefore,

$$\begin{aligned}
(1.4) &= -M^{-1} \frac{\vec{k} - M \cdot \vec{h}}{\|\vec{k}\|} \\
&= -M^{-1} \left(\frac{f(\vec{a} + \vec{h}) - f(\vec{h}) - M \cdot \vec{h}}{\|\vec{h}\|} \right) \cdot \left(\frac{\|\vec{h}\|}{\|\vec{k}\|} \right)
\end{aligned}$$

Since f is bi-Lipschitz on some E, $\frac{\|\vec{h}\|}{\|\vec{k}\|}$ is bounded by Lipschitz constant of f^{-1}. As $k \to 0$, we have that $h \to 0$. Therefore, by differentiability, we have that the other term goes to 0. Thus, ρ is differentiable. We still must show that $D\rho$ is continuous. Note that

$$D\rho(\vec{b}) = (Df(\rho(\vec{b})))^{-1}$$

Therefore, $D\rho$ is the composition of three continuous functions. Namely, $D\rho = \text{inversion} \circ Df \circ \rho$. Inversion is actually C^∞. To see this, use Cramer's rule. Thus, $\rho \in C^1$.

For $r > 1$, use induction. Consider the same composition. By our inductive hypothesis, we now have that ρ is C^{r-1}. From our hypothesis, Df is C^{r-1} and inversion is still C^∞. From homework, composition of C^r maps is C^r. Thus, $D\rho$ is C^{r-1} and $\rho \in C^r$. ■

Lecture 12: TBD

27 Sep. 02:30

1.5.4 Leibniz-notation, multivariable version

Suppose we have five quantities x, y, u, v, t such that any 3 are C^1 functions of the other two. One surface in $\mathbb{R}^5$, 10 graph interpretations. For affine example, choose "random" plane in $\mathbb{R}^5$. With probability 1, this is the case. Consider

$$f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} u \\ v \end{pmatrix}, \quad g : \begin{pmatrix} u \\ v \end{pmatrix} \mapsto t$$

Write $Df = \begin{bmatrix} D_1 u & D_2 u \\ D_1 v & D_2 v \end{bmatrix}$ as $\begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}$ Similarly,

$$Dg = \begin{bmatrix} D_1 t & D_2 t \end{bmatrix} = \begin{bmatrix} \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{bmatrix}$$

Thus,

$$\begin{aligned} \begin{bmatrix} \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} \end{bmatrix} &= D(g \circ f) \\ &= (Dg \circ f) \cdot Df \\ &= \begin{bmatrix} \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial t}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial t}{\partial v}\frac{\partial v}{\partial x} & \frac{\partial t}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial t}{\partial v}\frac{\partial v}{\partial y} \end{bmatrix} \end{aligned}$$

Biggest Danger: $\frac{\partial t}{\partial u}$ does not specify that v being fixed. Some subjects, a fix has been imposed. Use $\left(\frac{\partial t}{\partial u}\right)_v$.

Example. $u = x + y$, $v = x - y = 2x - u$. Then

$$\left(\frac{\partial v}{\partial x}\right)_y = 1 \neq 2 = \left(\frac{\partial v}{\partial x}\right)_u$$

1.6 Implicit Function Theorem

Start with affine warm-up. Given $M \in \operatorname{Mat}(n, k+n)$, $\vec{c} \in \mathbb{R}^n$. Let $E = \{\vec{v} \in \mathbb{R}^{k+n} \mid M\vec{v} = \vec{c}\}$. Let $\widehat{M}$ be the map $\vec{v} \mapsto M\vec{v}$. E could be empty (if $\vec{c} \neq 0$) or non-empty affine set $E = \vec{v}_{\text{part}} + \ker \widehat{M}$ with $\dim E = \dim \ker \widehat{M} = k + n - \operatorname{rank} M$. Therefore, $k \leq \dim E \leq k + n$.

Problem 1.6.1. Can we write $E = \operatorname{Graph}(g)$ for some g? (g must be affine by definition)

Answer. Suppose $\dim E = k$. Then we can write for some $X \in \mathbb{R}^{n \times k}$ and $Y \in \mathbb{R}^{n \times n}$,

$$M = \begin{bmatrix} X & Y \end{bmatrix},$$

so that $M \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} = X\vec{x} + Y\vec{y}$.

Case 1: Y not invertible.

Then $\exists \vec{w} \in \mathbb{R}^n \setminus \{\vec{0}\}$ such that $Y\vec{w} = \vec{0}$. Therefore, if $\begin{pmatrix} \vec{x} \\ \vec{w} \end{pmatrix} \in E$ then $\begin{pmatrix} \vec{x} \\ \vec{y} + \vec{w} \end{pmatrix} \in E$. Thus, E is not a graph.

Case 2: Y invertible.

Consider $M\begin{pmatrix}\vec{x}\\ \vec{y}\end{pmatrix} = \vec{c}$. Then

$$Y^{-1}X\vec{x} + \vec{y} = Y^{-1}\vec{c}$$

and

$$\vec{y} = \underbrace{-Y^{-1}X\vec{x} + Y^{-1}\vec{c}}_{=:T\vec{x}},$$

where T is affine and $E = \operatorname{Graph} T$.

Let $\mathrm{Proj}_{(k+1,\dots,n+k)} = \begin{bmatrix}0 & I\end{bmatrix}$ i.e. $\begin{pmatrix}\vec{x}\\ \vec{y}\end{pmatrix} \mapsto \vec{y}$, so that $Y = \mathrm{Proj}_{(k+1,\dots,n+k)} M$. Then E is a graph if and only if Y is invertible. ⊛

Note. Previous answer chose $x_1, \dots, x_k$ as independent variables. If Y is not invertible then we can still try with a different set of $x_{i_1}, \dots, x_{i_k}$. In particular, if M is rank k then we can permute the k independent columns to the front. This matrix $\widehat{p} : \mathbb{R}^{k+n} \to \mathbb{R}^{k+n}$ can be constructed as

$$\hat{p} = \begin{bmatrix}(\vec{e}_{i_1})^T\\ \vdots\\ (\vec{e}_{i_{k+n}})^T\end{bmatrix}$$

To summarize, $\dim E = k \Leftrightarrow \operatorname{rank} M = n \Leftrightarrow$ some $n \times n$ submatrix is invertible $\Leftrightarrow$ E is permuted graph of affine function (k independent variables).

Lecture 13: TBD

From handout, consider 29 Sep. 02:30

$$\{y^3 + xy + 1 = 0\}$$

- Can view this as a graph of function of $y \in \mathbb{R} \setminus \{0\}$ i.e. $x = -y^2 - \frac{1}{y}$.
- Can view as 3 graphs over x (plus one pt).
- Algebraically get 3 formula
  - One $\mathbb{R}$-valued,
  - Two $\mathbb{C}$-valued.

Theorem 1.6.1 (Implicit Function Theorem). Given $f \in C^r(A^{\text{osso } \mathbb{R}^{k+n}}, \mathbb{R}^n)$, $\begin{pmatrix}\vec{a}\\ \vec{b}\end{pmatrix} \in E := f^{-1}(\vec{0})$, $\frac{\partial f}{\partial y}\begin{pmatrix}\vec{a}\\ \vec{b}\end{pmatrix}$ invertible. Then there exists a neighborhood U of $\begin{pmatrix}\vec{a}\\ \vec{b}\end{pmatrix}$ such that $E \cap U = \operatorname{Graph}(g)$, $g \in C^r(B, \mathbb{R}^n)$, $a \in B^{\text{open}} \subset \mathbb{R}^k$.

Remark. $Df = \begin{bmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}\end{bmatrix}$ with $\frac{\partial f}{\partial x} \in \mathbb{R}^{n\times k}$ and $\frac{\partial f}{\partial y} \in \mathbb{R}^{n\times n}$. $\vec{x}$ is "independent" variables and $\vec{y}$ is "dependent" variables.

Notation. Write general point in $\mathbb{R}^{k+n}$ as $\begin{pmatrix}\vec{x}\\ \vec{y}\end{pmatrix}$ for $\vec{x} \in \mathbb{R}^k$ and $\vec{y} \in \mathbb{R}^n$.

Example. $f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto y^5 + xy + 1$. $Df = \begin{bmatrix} y & 5y^4 + x \end{bmatrix}$. $E := f^{-1}(0)$ is locally $y = g(x)$ except possibly when

$$\begin{cases} 5y^4 + x = 0 \\ y^5 + xy + 1 = 0 \end{cases} \Rightarrow x = \frac{-5}{2^{8/5}},\ y = \frac{1}{2^{2/5}}.$$

Lecture 14: TBD

Back to $y^5 + xy + z = 0$, order variables as (x, z, y). Then 2 Oct. 02:30

$$Df = \begin{bmatrix} y & 1 & 5y^4 + x \end{bmatrix}.$$

We can use implicit function theorem when $5y^4 + x \neq 0$. To find a "folding curve" we need to solve the system

$$\psi(x, z, y) = \begin{pmatrix} y^5 + xy + z \\ 5y^4 + x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Then

$$D\psi = \begin{bmatrix} y & 1 & 5y^4 + x \\ 1 & 0 & 20y^5 \end{bmatrix}$$

Note that the submatrix $\begin{bmatrix} y & 1 \\ 1 & 0 \end{bmatrix}$ is invertible. Therefore, we can choose (x, z) dependent and y independent. However, we can actually solve

$$\psi(x, z, y) = \begin{pmatrix} y^5 + xy + z \\ 5y^4 + x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} x &= -5y^4 \\ z &= 4y^5 \end{cases}$$

Therefore, $x = -\frac{5z^{4/5}}{2^{8/5}}$. Can use implicit function theorem to determine number of y-values for each (x, z).

1.6.1 Implicit Differentiation

Math 395 Approach: Given $f : \mathbb{R}^2 \to \mathbb{R}^2$, looking at Graph$(g)$ sitting in $f^{-1}(0)$. Let $\tilde{g}(x) = \begin{pmatrix} x \\ g(x) \end{pmatrix}$, so that $f \circ \tilde{g} = 0$. By chain rule,

$$Df(\tilde{g}(x)) \cdot \tilde{g}'(x) = 0$$

Therefore,

$$\begin{bmatrix} \frac{\partial f}{\partial x}\begin{pmatrix} x \\ g(x) \end{pmatrix} & \frac{\partial f}{\partial y}\begin{pmatrix} x \\ g(x) \end{pmatrix} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ g'(x) \end{bmatrix} = 0 \Leftrightarrow \frac{\partial f}{\partial x}\begin{pmatrix} x \\ g(x) \end{pmatrix} + \frac{\partial f}{\partial y}\begin{pmatrix} x \\ g(x) \end{pmatrix} g'(x) = 0$$

so $g'(x) = -\frac{\frac{\partial f}{\partial x}\begin{pmatrix} x \\ g(x) \end{pmatrix}}{\frac{\partial f}{\partial y}\begin{pmatrix} x \\ g(x) \end{pmatrix}}$.

In general dimension, let $g(\vec{x}) = \begin{pmatrix} \vec{x} \\ g(\vec{x}) \end{pmatrix}$. Once again

$$(Df \circ \tilde{g}) \cdot D\tilde{g} = 0$$

Therefore,

$$\begin{bmatrix} \frac{\partial f}{\partial \vec{x}} \circ \tilde{g} & \frac{\partial f}{\partial \vec{y}} \circ \tilde{g} \end{bmatrix} \cdot \begin{bmatrix} I \\ Dg \end{bmatrix} = 0$$

Then

$$\underbrace{\frac{\partial f}{\partial \vec{x}} \circ \tilde{g}}_{n \times k} + \underbrace{\left(\frac{\partial f}{\partial \tilde{y}} \circ \tilde{g}\right)}_{n \times n} \cdot \underbrace{Dg}_{n \times k}$$

Thus,

$$Dg = -\left(\frac{\partial f}{\partial \vec{y}}\right)^{-1} \circ \tilde{g} \cdot \frac{\partial f}{\partial \vec{x}} \circ \tilde{g}$$

Math 295 Approach: Back to $y^5 + xy + 1 = 0$. Assuming that $g(x) = y$ differentiable function of x, have

$$5y^4 \frac{dy}{dx} + y + x\frac{dy}{dx} = 0$$

Thus,

$$\frac{dy}{dx} = -\frac{y}{5y^4 + x}$$

then at $(0, -1)$, $\frac{dy}{dx} = \frac{1}{5}$. Additionally, we have the following ODE

$$g'(x) = -\frac{g(x)}{5g^4(x) + x}$$

1.6.2 Proof of Implicit Function Theorem

Proof. Let $F\begin{pmatrix}\vec{x}\\ \vec{y}\end{pmatrix} = \begin{pmatrix}\vec{x}\\ f\begin{pmatrix}\vec{x}\\ \vec{y}\end{pmatrix}\end{pmatrix}$. $F = C^r(A, \mathbb{R}^{k+n})$ and

$$DF = \begin{bmatrix} I_{k\times k} & 0_{k\times n} \\ \frac{\partial f}{\partial \vec{x}} & \frac{\partial f}{\partial \vec{y}} \end{bmatrix}$$

From Munkres 2.6,

$$\det DF = \det \frac{\partial f}{\partial \vec{y}} \neq 0.$$

So $DF\begin{pmatrix}\vec{a}\\ \vec{b}\end{pmatrix}$ is invertible. As an exercise, check

$$DF\begin{pmatrix}\vec{a}\\ \vec{b}\end{pmatrix}^{-1} = \begin{bmatrix} I & 0 \\ -\left(\frac{\partial f}{\partial \vec{y}}\right)^{-1} \cdot \frac{\partial f}{\partial \vec{x}} & \left(\frac{\partial f}{\partial \vec{y}}\right)^{-1} \end{bmatrix}$$

By inverse function theorem, F has local C^r inverse

$$F^{-1}\begin{pmatrix}\vec{x}\\ \vec{y}\end{pmatrix} = \begin{pmatrix}\vec{x}\\ h\begin{pmatrix}\vec{x}\\ \vec{y}\end{pmatrix}\end{pmatrix},$$

where $h \in C^r$. Near $\begin{pmatrix}\vec{a}\\ \vec{b}\end{pmatrix}$ have

$$\begin{aligned}\begin{pmatrix}\vec{x}\\ \vec{y}\end{pmatrix} \in E &\Leftrightarrow f\begin{pmatrix}\vec{x}\\ \vec{y}\end{pmatrix} = \vec{0}\\ &\Leftrightarrow F\begin{pmatrix}\vec{x}\\ \vec{y}\end{pmatrix} = \begin{pmatrix}\vec{x}\\ \vec{0}\end{pmatrix}\\ &\Leftrightarrow \begin{pmatrix}\vec{x}\\ \vec{y}\end{pmatrix} = F^{-1}\begin{pmatrix}\vec{x}\\ \vec{0}\end{pmatrix} = \begin{pmatrix}\vec{x}\\ h\begin{pmatrix}\vec{x}\\ \vec{0}\end{pmatrix}\end{pmatrix}\\ &\Leftrightarrow \vec{y} = h\begin{pmatrix}\vec{x}\\ \vec{0}\end{pmatrix}\end{aligned}$$

■

Remark. A few remarks:

(1) F is a C^r-diffeomorphism (or "C^r change of coordinates")

- Take $\vec{y} = h\begin{pmatrix}\vec{x}\\ \vec{x}\end{pmatrix}$, then if you apply F you "straighten" out $\vec{y} = \vec{c}$ in some way?

(2) E is locally a coordinate-permuted graph if some $n\times n$ submatrix is invertible or rank $Df\begin{pmatrix}\vec{a}\\ \vec{b}\end{pmatrix} = n$ (Df has maximal rank).

Chapter 2

Constrained Optimization

2.1 Motivation

Situation 1: No constraints

Suppose $h \in C^1(\Omega^{\text{osso } \mathbb{R}^n}, \mathbb{R})$. h has max/min of $\vec{p} \in \Omega$. Consider $\gamma \in C^1$ (osso $\mathbb{R}$, Ω) with $\gamma(0) = \vec{p}$. Then $h \circ \gamma$ has max/min at $t = 0$, so

$$0 = (h \circ \gamma)'(0) = Dh(\gamma(0)) \cdot \gamma'(0)$$

Conclude that $Dh(\vec{p}) = 0$ i.e.

$$0 = D_1 h(\vec{p}) = D_2 h(\vec{p}) = \cdots = D_n h(\vec{p})$$

n equations and n unknowns.

Situation 1': As above but h has local max/min.

Same conclusion.

Situation 2: Constraints

Given $f \in C^1(\Omega^{\text{osso } \mathbb{R}^{k+n}}, \mathbb{R}^n)$, $\vec{p} \in E := f^{-1}(0)$, $h \in C^1(\Omega, \mathbb{R})$, $h|_E$ has local max/min at $\vec{p}$. Consider $\gamma \in C^1(\text{osso } \mathbb{R}, E)$ with $\gamma(0) = \vec{p}$ i.e. $\gamma \in C^1(\text{ooso } \mathbb{R}, \mathbb{R}^{n+k})$ with values in E.

- $h \circ \gamma$ has local max/min at $t = 0$, so $0 = (h \circ \gamma)'(0) = Dh(\vec{p}) \cdot \gamma'(0)$.

What do we know about $\gamma'(0)$? Note

$$\begin{aligned} f \circ \gamma &\equiv 0 \\ Df(\gamma(t)) \cdot \gamma'(t) &= 0 \\ Df(\vec{p}) \cdot \gamma'(0) &= 0 \end{aligned}$$

Therefore, $\gamma'(0) \in \ker Df(\vec{p})$.

Lemma 2.1.1. If $Df(\vec{p})$ has maximal rank, then there are no other constraints on $\gamma'(0)$.

Proof. HW6 1. Use Implicit Function Theorem. ■

Continuing to assume $Df(\vec{p})$ has maximal rank and combining, we have

$$Dh(\vec{p}) \in (\ker Df(\vec{p}))^\perp = \left((\text{rowspace}\, Df(\vec{p}))^\perp\right)^\perp = \text{rowspace}\, Df(\vec{p}) = \text{span}\{Df_1(\vec{p}), \ldots, Df_n(\vec{p})\}$$

so

$$Dh(\vec{p}) = \lambda_1 Df_1(\vec{p}) + \cdots + \lambda_n Df_n(\vec{p}).$$

Since $Df(\vec{p})$ has rank n, Df_i's are linearly independent and the λ_i's are uniquely determined. These λ_i's are called *Lagrange multipliers*. This gives us the unknowns $(\vec{p},\ \lambda) \in \mathbb{R}^{k+2n}$ and the system

$$\begin{aligned} f(\vec{p}) &= 0 \\ Dh(\vec{p}) &= \sum_{i=1}^{n} \lambda_i Df_i(\vec{p}) \end{aligned}$$

or $k + 2n$ constraints. Therefore, we have the same number of constraints as unknowns. If $\det Df(\vec{p}) \neq 0$ can apply Inverse Function Theorem.

Globally, $\emptyset \neq K^{\text{compact}} \subset \mathbb{R}^n$, $h : K \to \mathbb{R}$ continuous. By EVT, h has max/min on K. What points need to be checked?

(1) $\vec{p} \in \operatorname{Int} K$ then $Dh(\vec{p}) = 0$,

(2) $\vec{p} \in \operatorname{Int} K$, h not differentiable at $\vec{p}$.

(3) $\vec{p} \in \operatorname{Bd} K$.

Lecture 15: TBD

4 Oct. 02:30

Example. max/min $h(x, y) = x^4 + y^6$ on $K = \{(x, y) \mid x^2 + y^2 \leq 1\}$.

Min at $\vec{0}$. $Dh(x, y) = \begin{bmatrix} 4x^3 & 6y^5 \end{bmatrix}$. Since $Dh = 0$ only at $\vec{0}$, we attain a maximum on the boundary. Consider $f\begin{pmatrix} x \\ y \end{pmatrix}$ so that $f^{-1}(0) = \operatorname{Bd} K$. Then $Df = \begin{bmatrix} 2x & 2y \end{bmatrix}$. This yields the following system

$$\begin{aligned} x^2 + y^2 &= 1 \\ 4x^3 &= \lambda 2x \\ 6y^5 &= \lambda 2y \end{aligned}$$

Solving yields, $x = \pm\sqrt{\lambda/2}$, $y = \sqrt[4]{\lambda/2}$ with $\lambda = \frac{2}{3}(4 - \sqrt{z})$. So $\max h = 1$.

Exercise. Replace $x^2 + y^2 \leq 1$ with $x^6 + y^6 \leq 1$.

Lemma 2.1.2. Given $X^{\text{closed}} \subset \mathbb{R}^n$, $\vec{p} \in \mathbb{R}^n$. Then $\vec{x} \mapsto \|x - p\|^2$ has a minimum on X.

Proof. HW5 6. ∎

Example. What point(s) of $xyz = 1$ lies closest to $\vec{0}$?

Consider $f(x, y, z) = xyz - 1$, $h(x, y, z) = x^2 + y^2 + z^2$. Then we look for solutions to $Dh = \lambda Df$,

$$\begin{bmatrix} 2x & 2y & 2z \end{bmatrix} = \lambda \begin{bmatrix} yz & xz & xy \end{bmatrix}$$

Additionally, we have that $xyz = 1$. Can simplify to

$$\begin{aligned} 2x^2 = 2y^2 = 2z^2 &= \lambda \\ xyz &= 1 \end{aligned}$$

Then min dist is $\sqrt{3}$.

Exercise. Distance from $\{x^a y^b z^c = 1\}$ to $\vec{0}$.

Example. Let $B \in \mathrm{Mat}(n, n)$ symmetric. Let $h(\vec{x}) = \vec{x}^T B \vec{x}$ and $K = \|\vec{x}\| = 1$.

Consider $f(x) = \|\vec{x}\|^2 - 1$. Check: $\operatorname{rank} Df(\vec{x}) = 1$ on $\|\vec{x}\|^2 = 1$. $Df(\vec{x}) = 2\vec{x}^T$. max exists, occurs at solution of $Dh = \lambda Df$.

Claim. $Dh(\vec{x}) = 2\vec{x}^T B$.

Proof (1). Note that

$$h(\vec{x}) = \sum_{j,k} b_{j,k} x_j x_k$$

Therefore,

$$D_m h(\vec{x}) = \sum_k b_{m,k} x_k + \sum_j {}_{j,m} x_j = 2 \sum_j b_{j,m} x_j = (2\vec{x}^T B)_m$$

■

Proof (2). Note that

$$\begin{aligned} Dh(\vec{x}) \cdot \vec{u} &= h'(\vec{x}; \vec{u}) \\ &= \vec{u}^T B \vec{x} + \vec{x}^T B \vec{u} \\ &= 2\vec{x}^T B \vec{u} \end{aligned}$$

■

Therefore, we need that

$$Dh(\vec{x}) = \lambda Df(\vec{x}) \quad \Leftrightarrow \quad 2\vec{x}^T B = 2\lambda \vec{x}^T$$

i.e. $B\vec{x} = \lambda \vec{x}$. Note

$$h(\vec{x}) = \vec{x}^T B \vec{x} = \lambda$$

i.e. max over sphere.

Lecture 16: TBD

6 Oct. 02:30

Example (Followup). Let $\vec{x}_1$ be an eigenvector with eigenvalue μ_1. Instead of max over sphere, consider max h over $S^{n-1} \cap \{\vec{x}_1\}^\perp$ with

$$f(\vec{x}) = \begin{pmatrix} \|\vec{x}\|^2 - 1 \\ \vec{x}_1^T \cdot \vec{x} \end{pmatrix} \coloneqq \begin{pmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \end{pmatrix}.$$

Need $Dh(\vec{x}) = \lambda_1 Df_1(\vec{x}) + \lambda_2 Df_2(\vec{x})$ i.e.

$$\begin{aligned} &\vec{x}^T \cdot \vec{x} = 1 \\ &\vec{x}_1^T \cdot \vec{x} = 0 \\ &2\vec{x}^T B = 2\lambda_1 \vec{x}^T + \lambda_2 \vec{x}_1^T \end{aligned}$$

Therefore,

$$2\vec{x}^T B \vec{x}_1 = \lambda_2 = 0$$

This gives us that

$$\vec{x}^T B = \lambda_1 \vec{x}^T \quad \Leftrightarrow \quad B\vec{x} = \lambda_1 \vec{x}$$

Therefore, $\lambda_2 \coloneqq \mu_2$ is an eigenvalue with eigenvector $\vec{x}_2 \in \{\vec{x}_1\}^\perp$.

Theorem 2.1.1. B symmetric (real)-matrix emits an orthonormal basis of eigenvectors $\vec{x}_1, \ldots, \vec{x}_n$ with real eigenvalues $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$.

Theorem 2.1.2. For constant function f, $\vec{p} \in \mathcal{U}^{\text{osso } \mathbb{R}^{k+n}}$, $f \in C^1(\mathcal{U}, \mathbb{R}^n)$, $\operatorname{rank} Df(\vec{p}) = n$, $E = f^{-1}(\vec{0})$. For objective function $h \in C^1(\mathcal{U}, \mathbb{R})$, $h|_E$ has a local maximum at $\vec{p}$ and

$$Dh(\vec{p}) \in \operatorname{rowspace} Df(\vec{p}) \quad \Leftrightarrow \quad Dh(\vec{p}) = \lambda_1 Df_1(\vec{p}) + \cdots + \lambda_n Df_n(\vec{p})$$

Lecture 17: TBD

6 Oct. 02:30

Definition. $h(\vec{x}) = \vec{x}^T B \vec{x}$. Then

$$h(c_1\vec{x}_1 + \cdots + c_n\vec{x}_n) = c_1^2\mu_1 + \cdots + c_n^2\mu_n$$

Definition 2.1.1. All $\mu_j \geq 0$ if and only if $\vec{x}^T B \vec{x} \geq 0$ for all $\vec{x} \in \mathbb{R}^n$ if and only if $B \geq 0$ *positive semi definite.*

Definition 2.1.2. All $\mu_K > 0$ if and only if $\vec{x}^T B \vec{x} > 0$ for all $\vec{x} \in \mathbb{R}^n \setminus \{0\}$ if and only if $B > 0$ *positive definite.*

Definition 2.1.3. $B < (\leq) 0$ if and only if $(-B) > (\geq) 0$ *negative (semi-)definite.*:

Theorem 2.1.3. Given $\Omega^{\text{convex open}} \subset \mathbb{R}^n$, $f \in C^2(\Omega, \mathbb{R})$,

$$Hf(\vec{x}) \coloneqq (D_j D_k f(\vec{x}))_{j,k} \geq 0 \ \forall \vec{x} \in \Omega.$$

This the *Hessian* of f at $\vec{x}$. $Hf(\vec{x}) \in \operatorname{Symm}(n) = \{M \in \operatorname{Mat}(n,n) \mid M^T = M\}$. If $Df(\vec{x}_0) = 0$ then $f(\vec{x}) \geq f(\vec{x}_0)$ for all $\vec{x} \in \Omega$

Proof. Let $\phi(t) = (1-t)\vec{x}_0 + t\vec{x}$, so that $\phi'(t) = \vec{x} - \vec{x}_0 \coloneqq \vec{a}$. Then

$$(f \circ \phi)'(t) = Df(\phi(t)) \cdot \vec{a} = \sum_j D_j f(\phi(t)) \cdot a_j.$$

Note that

$$(f \circ \phi)'(0) = Df(\vec{x}_0) \cdot \vec{a} = 0.$$

Similarly,

$$\begin{aligned}
(f \circ \phi)''(t) &= \sum_j a_j D D_j f(\phi(t)) \cdot \phi'(t) \\
&= \sum_{j,k} a_j a_k D_j D_k f(\phi(t)) \\
&= \vec{a}^T H f(\phi(t)) \vec{a} \geq 0.
\end{aligned}$$

By Theorem F.1 in the supplement,

$$f(\vec{x}) = (f \circ \phi)(1) \geq (f \circ \phi)(0) = f(\vec{x}_0).$$

■

Addendun: If also $Hf(\vec{x}) > 0$, $\forall x \in \Omega$ then $f(\vec{x}) > f(\vec{x}_0)$ for $\vec{x} \in \Omega \setminus \{\vec{x}_0\}$.

Theorem 2.1.4. Let $\mathrm{Pos}(n) :=$ set of positive definite matrices. Then $\mathrm{Pos}(n)$ is open in $\mathrm{Symm}(n)$.

Proof. Proof in supplement. Sufficient to consider $\vec{a}$ such that $\|\vec{a}\| = 1$. Use compactness of unit sphere to apply EVT. ■

2.2 2nd Derivative Test

Consider $f \in C^2(\text{osso } \mathbb{R}^n, \mathbb{R})$ and $\vec{x}_0 \in \mathrm{dom}(f)$ with $Df(\vec{x}_0) = 0$. $\vec{x}_0$ is a *critical point* for f.

(a) $Hf(\vec{x}_0) > 0$ then $Hf(\vec{x}) > 0$ for $\vec{x} \in \mathcal{U}(\vec{x}_0, \delta)$. Therefore, f has strict local min at $\vec{x}_0$.

(b) $Hf(\vec{x}_0) \not\geq 0$ f has strict local max along some line through $\vec{x}_0$. f does not have a local min at $\vec{x}_0$.

(c) $Hf(\vec{x}_0) < 0$ then f has strict local max at $\vec{x}_0$.

(d) $Hf(\vec{x}_0) \not\leq 0$ then f does not have a local max.

(e) $Hf(\vec{x}_0) \not\geq 0, \not\leq 0$ then f does not have a local max or min at $\vec{x}_0$.

Corollary 2.2.1. Given $f \in C^2(\Omega^{\text{convex oss } \mathbb{R}^n}, \mathbb{R})$, $Hf \geq 0$ on Ω then

$$f(\vec{x}) \geq f(\vec{x}_0) + Df(\vec{x}_0)(\vec{x} - \vec{x}_0) \quad \forall \vec{x}_0, \vec{x} \in \Omega$$

Proof. Let $g(x) = f(\vec{x}) - Df(\vec{x}_0)(\vec{x} - \vec{x}_0)$, so that $Dg(\vec{x}) = Df(\vec{x}) - Df(\vec{x}_0)$. Then $Dg(\vec{x}_0) = 0$. Since $Hg(\vec{x}) = Hf(\vec{x})$, from a previous theorem, $g(\vec{x}) \geq g(\vec{x}_0) = f(\vec{x}_0)$ as desired. ■

Definition 2.2.1. For $f : \Omega \to \mathbb{R}$ define the *epigraph* of f as the set $\{(\vec{x}, y) \in \Omega \times \mathbb{R} \mid y \geq f(x)\}$.

Definition 2.2.2. For $f : \Omega \to \mathbb{R}$ define the *hypograph* of f as the set $\{(\vec{x}, y) \in \Omega \times \mathbb{R} \mid y \leq f(x)\}$.

Corollary 2.2.2. Given $f \in C^2(\Omega^{\text{convex oss } \mathbb{R}^n}, \mathbb{R})$, $Hf \geq 0$ on Ω then

$$\mathrm{epigraph}(f) = \bigcap_{\vec{x}_0 \in \Omega} \{(\vec{x}, y) \in \Omega \times \mathbb{R} \mid y \geq f(\vec{x}_0) + Df(\vec{x}_0)(\vec{x} - \vec{x}_0)\}$$

Proof. $(\vec{x}, y) \in \mathrm{epigraph}(f) \Rightarrow (\vec{x}, y) \in RHS$. Now suppose that $(\vec{x}, y) \in RHS$. Then if we choose $\vec{x}_0 = \vec{x}$ then $y \geq f(\vec{x}) \Rightarrow (\vec{x}, y) \in \mathrm{epigraph}(f)$. ■

Remark. We call $\{(\vec{x}, y) \in \Omega \times \mathbb{R} \mid y \geq f(\vec{x}_0) + Df(\vec{x}_0)(\vec{x} - \vec{x}_0)\}$ a *half-space*. Since this is clearly convex, we have that $\mathrm{epigraph}(f)$ is convex.

Definition 2.2.3. Given $\Omega^{\text{convex}} \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$. f is convex if and only if $\mathrm{epigraph}(f)$ is convex.

Midterm stuff

This is the cutoff for midterm material.

- The midterm will take place on Thursday 10/19 in WEIS260 7-9pm.
- Can bring in some notes (TBD).

Chapter 3

Integration

Lecture 19: TBD

$f : Q \to \mathbb{R}$ bounded. Define 13 Oct. 02:30

$$\operatorname{osc}(f, \vec{a}) \coloneqq \lim_{\delta \to 0} \left\{ \sup_{\mathcal{U}(a,\delta)\cap Q} f - \inf_{\mathcal{U}(a,\delta)\cap Q} f \right\}$$

$\{\vec{a} \in Q \mid \operatorname{osc}(f, \vec{a}) < \epsilon\}$ is (relatively) open. Let $\mathcal{D}_k \coloneqq \left\{\vec{a} \in Q \mid \operatorname{osc}(f, \vec{a}) \geq \frac{1}{k}\right\}$. Then

$$\mathcal{D} = \bigcup_{k=1}^{\infty} \mathcal{D}_k = \{\vec{a} \in Q \mid f \text{ not at } \vec{a}\}$$

may not be closed.

Example. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \begin{cases} x & \text{if } \frac{1}{x} \in \mathbb{N} \\ 0 & \text{if else} \end{cases}$. Then $\mathcal{D} = \{\frac{1}{n} \mid n \in \mathbb{N}\}$ is not closed.

Theorem 3.0.1. For $f : Q^{\text{box}} \to [-M, M]$, the following are equivalent

(1) f is (Riemman)-integrable on Q

(2) For $\epsilon > 0$ there is a partition P with $\mathcal{U}(f, p) < L(f, p) + \epsilon$.

(3) For $\epsilon > 0$, $k \in \mathbb{N}$ can write

$$\mathcal{D}_k \subset R_1 \cup \cdots \cup R_j$$

with $v(R_i) + \cdots + v(R_j) < \epsilon$, where v is volume

(4) For $\epsilon > 0$, can write $\mathcal{D}_k \subset \bigcup_{p=1}^{\infty} R_p$ with $\sum_{p=1}^{\infty} v(R_p) < \epsilon$.

(5) For $\epsilon > 0$ can write $\mathcal{D} \subset \bigcup_{p=1}^{\infty} \operatorname{relint}(R_p^{\text{box}})$ with $\sum_{p=1}^{\infty} v(R_p) < \epsilon$.

(1) $\Leftrightarrow$ (2). Done. ■

(2) $\Rightarrow$ (3). Pick P such that

$$\sum_{R \text{ defined by } P} (\sup_R f - \inf_R f) v(R) < \frac{\epsilon}{k}$$

Let $R_1, \dots, R_\ell$ be boxed defined by P whose interior meets $\mathcal{D}_k$. Then

$$\frac{1}{k}\sum_{p=1}^{\ell} v(R_p) \le \sum_{p=1}^{\ell}(\sup_{R_p} f - \inf_{R_p} f)v(R_p) < \frac{\epsilon}{k}$$

so $\sum_{p=1}^{\ell} v(R_p) < \epsilon$. But $\mathcal{D}_k \subset \bigcup_{p=1}^{\ell} R_p \cup \bigcup_{p=1}^{m} \operatorname{Bd}\tilde{R}_i$, where $\tilde{R}_i$ are volume zero boxes. Thus, sum $\le \epsilon$. ■

Proof of (3) ⇒ (4). Cover $\mathcal{D}_k$ with finitely many boxes with volume sum $\le \frac{\epsilon}{2^k}$. Combine. ■

Lecture 20: TBD

18 Oct. 02:30

Lemma 3.0.1. Given $R^{\text{box}} \subset Q$, $y > v(R)$, then there exists some box $\tilde{R}$ such that $R \subset \operatorname{relint}(\tilde{R})$ and $v(\tilde{R}) < y$.

Proof of (4) ⇒ (5). Pick $\mathcal{D} \subset \bigcup_{p=1}^{\infty} R_p$ with $\sum v(R_p) < \epsilon/4$. Pick $R_p \subset \operatorname{relint}(\tilde{R}_p) \subset \tilde{R}_p \subset Q$ with $v(\tilde{R}_p) < 2v(R_p)$ if $v(R_p) \ne 0$. Otherwise, $v(\tilde{R}_p) < \frac{\epsilon}{2^{p+1}}$. ■

Proof of (5) ⇒ (2). Given $\epsilon > 0$ let $\tilde{\epsilon} = \frac{\epsilon}{3(M+v(Q))}$. Write $\mathcal{D} \subset_{p=1}^{\infty} \operatorname{relint}(R_p)$ with $\sum v(R_p) < \tilde{\epsilon}$.

$$Q = \bigcup_{p=1}^{\infty} \operatorname{relint}(R_p) \cup \bigcup_{\vec{a} \in Q \setminus \mathcal{D}} \operatorname{relint}(Q_{\vec{a}})$$

Pick finite subcover $Q \subset \operatorname{relint}(R_{p_1}) \cup \dots \cup \operatorname{relint}(R_{p_s}) \cup \operatorname{relint}(Q_{\vec{a_1}}) \cup \dots \cup \operatorname{relint}(Q_{\vec{a_n}})$. Choose partition P such that each sub-box T defined by P is contained by some R_{p_j} or some $Q_{\vec{a}_k}$.Then

$$\begin{aligned} \mathcal{U}(f,p) - L(f,p) &= \sum(\sup f - \inf f) \cdot v(R) + \sum(\sup f - \inf f) \cdot v(R) \\ &\le 2M\tilde{\epsilon} + 2v(Q))\tilde{\epsilon} \\ &< \epsilon \end{aligned}$$

■

(4) leads us to consider: For $E \subset \mathbb{R}^n$ let $m^*(E) \coloneqq \inf\{\sum_{j=1}^{\infty} v(Q_j) \mid E \subset \bigcup_{j=1}^{\infty} Q_j^{\text{box}}\}$.

3.0.1 Crossover Episode!

Definition 3.0.1. E has outer measure 0 ⇔ $m^*(E) = 0$ ⇔ E has measure 0.

Corollary 3.0.1. Bounded $f : Q^{\text{box}} \to \mathbb{R}$ integrable if and only if $m^*(\mathcal{D}_f) = 0$.

Problem 3.0.1. Why do we insist on countable coverings?

Answer. To avoid every E having measure 0. ⊛

Problem 3.0.2. Why do we allow infinite coverings?

Answer. Finite option is not ridiculous. Called **Jordan measure** or **Jordan content**. ⊛

Definition. $m^{*,J} = \inf\{\sum_{j=1}^{k} v(Q_j) \mid E \subset \bigcup_{j=1}^{k} Q_j^{\text{box}}\}$.

Note. $m^*(E) \le m^{*,J}(E)$.

Lemma 3.0.2. $m^*(E) = \inf\left\{\sum_{j=1}^\infty v(Q_j) \mid \bigcup_{j=1}^\infty \operatorname{Int} Q_j^{\text{box}}\right\}$.

Proof. Restrictions only increase inf's. Need to show no actualy increase occurs.

Case 1: $m^+(E) = +\infty$

Trivial.

Case 2: $m^+(E) < +\infty$

Q_j's converge. Pick $\tilde{Q}_j \supset \operatorname{Int} \tilde{Q}_j \supset Q_j$, with $v(\tilde{Q}_j) < v(Q_j) + \frac{\epsilon}{2^j}$. Cover E with $\bigcup \tilde{Q}_j$. Note

$$\sum v(\tilde{Q}_j) < \sum v(Q_j) + \epsilon$$

Thus, infs are the same.

■

Proposition 3.0.1. Suppose $K^{\text{compact}} \subset \mathbb{R}^n$. Then $m^{*,J}(E) = m^*(E)$.

Proof. Need $m^*(K) \geq m^{*,J}(K)$. Pick boxes Q_j with $\bigcup_{j=1}^\infty \operatorname{Int} Q_j \supset K$. By compactness, some $\bigcup_{k=1}^M \operatorname{Int} Q_{j_k}$. Then

$$\sum_{j=1}^\infty v(Q_j) \geq \sum_{k=1}^M v(Q_{j_k}) \geq m^{*,J}(K)$$

Take infimum over choice of Q_{j_k}'s. ■

Problem 3.0.3. Do Lebesgue and Jordan disagree?

Answer. Yes. Consider $E = \mathbb{Q} \cap [0, 1]$. Then $m^*(E) = 0$. However, $m^{*,J}(E) = 1$. ⊛

Theorem 3.0.2 (Fubini's Theorem)**.** Consider $A^{\text{box}} \subset \mathbb{R}^k$, $B^{\text{box}} \subset \mathbb{R}^n$, $Q = A \times B$, $f : \mathbb{Q} \to \mathbb{R}$ bounded. Then

$$\underline{\int}_Q f \leq \underline{\int}_{\vec{x}\in A} \left(\underline{\int}_{\vec{y}\in B} f(\vec{x}, \vec{y}) \right) \leq \begin{cases} \underline{\int}_{\vec{x}\in A} \overline{\int}_{\vec{y}\in B} f(\vec{x}, \vec{y}) \\ \overline{\int}_{\vec{x}\in A} \underline{\int}_{\vec{y}\in B} f(\vec{x}, \vec{y}) \end{cases} \leq \overline{\int}_{\vec{x}\in A} \overline{\int}_{\vec{y}\in B} f(\vec{x}, \vec{y}) \leq \overline{\int}_Q f$$

Corollary 3.0.2. If f integrable on Q all are equal. $\int_{\vec{y}\in B} f(\vec{x}, \vec{y})$ may not exist.

Corollary 3.0.3. f continuous on $Q = I_1 \times \cdots \times I_n$ then

$$\int_Q f = \int_{x_1\in I_1} \cdots \int_{x_n\in I_n} f(x_1, \ldots, x_n)$$

Lecture 21: TBD

Recall: 20 Oct. 02:30

- $\underline{\int}_Q f \leq \overline{\int}_Q f$.
- $f \leq g$ on $Q \Rightarrow \underline{\int}_Q f \leq \underline{\int}_Q g$.

Example. $Q = [-1,1] \times [-1,1]$. $f = \mathbf{1}_{\{0\}\times\mathbb{Q}}$. $\mathcal{D}_f = \{0\} \times [-1,1]$ zero volume box. From last lecture, f integrable on Q.

$$\overline{\int}_{y\in[-1,1]} f(x,y) = \begin{cases} 0, & \text{if } x \neq 0 \\ 1, & \text{if } x = 0 \end{cases}, \quad \underline{\int}_{y\in[-1,1]} f(x,y) = 0$$

Thus,

$$\int_Q f = 0$$

Not guaranteed for y integrals to match.

Lecture 22: TBD

Misc Facts: For $f : Q^{\text{box}} \to \mathbb{R}$ bounded, 23 Oct. 02:30

(i) If $f^{-1}(0)$ dense then $\underline{\int} f \leq 0 \leq \overline{\int} f$.

(ii) $f^{-1}(0)$ dense and f is integrable then $\int f = 0$.

(iii) $f \geq 0$, $f(\vec{a}) > 0$, f continuous at $\vec{a}$ then $\underline{\int} f > 0$.

Consider $S^{\text{bdd}} \subset \mathbb{R}^n$, $f : S \mapsto \mathbb{R}$ bounded. Define

$$f_S(\vec{x}) = \begin{cases} f(\vec{x}), & \text{if } \vec{x} \in S \\ 0, & \text{if } \vec{x} \notin S \end{cases}$$

Then $\int_S f := \int_Q f_S$ for $Q^{\text{box}} \supset S$.

Proposition 3.0.2. The existence and value of the RHS do not depend on choice of Q.

Proof. Recall existence of $\int_Q f_S$ depends on $m^*(\mathcal{D}_{f_S}) = 0$. Suffices to show that $\int_{Q_1} f_S = \int_{Q_3} f_S$. P partition of Q_3. Refine P to P' such that Q_1 is finite sum of P'. Then

$$\begin{aligned} L(f_S, P) \leq L(f_S, P') &= \underbrace{\sum_{P', R \subset Q_1} (\inf_R f_S) \cdot v(R)}_{\leq \underline{\int}_{Q_1} f_S} + \underbrace{\sum_{P', R \subset Q_3 \setminus \text{relint}(Q_1)} (\inf_R f_S) \cdot v(R)}_{\leq 0} \\ &\leq \int_{Q_1} f_S \end{aligned}$$

Taking sup over choice of P gives $\int_{Q_3} f_S \leq \int_{Q_1} f_S$. For the reverse direction, repeat using $U(f_S, P)$. ■

Rules:

(a) f, g integrable over S then $\int_S (af + bg) = a \int_S f + b \int_S g$.

(b) f, g integrable on S, $f \leq g$ on S then $\int_S f \leq \int_S g$.

  (b') $\left| \int f \right| \leq \int |f|$.

(c) If $f \geq 0$ on S and $T \subset S$, then $\int_T f \leq \int_S f$.

(d) If f integrable on S_1 and on S_2 then f integrable on $S_1 \cup S_2$ and $S_1 \cap S_2$, where

$$\int_{S_1 \cup S_2} f = \int_{S_1} f + \int_{S_2} f - \int_{S_1 \cap S_2} f$$

Example. $S = \{(x_1, x_2, x_3) \mid x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_1+x_2+x_3 \leq 1\}$. Take $Q = [0,1]\times[0,1]\times[0,1]$. Check that $\int_S 1$ exists. HW8 study exercise 1. By Fubini,

$$\begin{aligned}
\int_S 1 &= \int_Q \mathbf{1}_S \\
&= \int_{0\leq x_1\leq 1}\int_{0\leq x_2\leq 1}\int_{0\leq x_3\leq 1} \mathbf{1}_S \\
&= \int_{0\leq x_1\leq 1}\int_{0\leq x_2\leq 1} \max\{0, 1-x_1-x_2\} \\
&= \int_{0\leq x_1\leq 1}\int_{0\leq x_2\leq 1-x_1} 1-x_1-x_2 \\
&= \int_{0\leq x_1\leq 1} (1-x_1)x_2 - \frac{x_2^2}{2}\Big|_{x_2=0}^{x_2=1-x_1} \\
&= \int_{0\leq x_1\leq 1} \frac{(1-x_1)^2}{2} \\
&= -\frac{(1-x_1)^3}{6}\Big|_{x_1=0}^{x_1=1} \\
&= \frac{1}{6}
\end{aligned}$$

Theorem 3.0.3. Given $S^{\text{bdd}} \subset \mathbb{R}^n$, $f : S \to \mathbb{R}$ bounded continuous. Define

$$E \coloneqq \{\vec{x_0} \in \operatorname{Bd} S \mid \text{not true that } \lim_{\vec{x}\to\vec{x_0}} f(\vec{x}) = 0\}.$$

with $m^*(E) = 0$. Then f is integrable on S.

Proof. $\mathcal{D}_{f_S} \subset E$ so $m^*(\mathcal{D}_{f_S}) = 0$. Thus, f_S is integrable. ∎

Lecture 23: TBD

25 Oct. 02:30

Corollary 3.0.4. Given $S^{\text{bdd}} \subset \mathbb{R}^n$, $f : S \to \mathbb{R}$ continuous bounded, $m^*(\operatorname{Bd} S) = 0$ then f integrable on S.

Definition 3.0.2. S is *rectifiable*

$$\begin{aligned}
&\Leftrightarrow \mathbf{1} \text{ is integrable on } S \\
&\Leftrightarrow \mathbf{1}_S \text{ is integrable on } Q^{\text{box}} \supset S \\
&\Leftrightarrow m^*(\operatorname{Bd} S) = 0 \\
&\Leftrightarrow m^{*,J}(\operatorname{Bd} S) = 0
\end{aligned}$$

Definition 3.0.3. $v(S) = \int_S \mathbf{1} = \mathbf{1}_S$.

S rectifiable, $A = \operatorname{Int} S$, then $\operatorname{Bd} A \subset \operatorname{Bd} S$. Therefore, $m^*(\operatorname{Bd}(A) \leq m^*(\operatorname{Bd}(S)) = 0$. Therefore A is rectifiable and $\mathbf{1}_S$, $\mathbf{1}_A$ integrable. $\mathbf{1}_{S\setminus A} = \mathbf{1}_S - \mathbf{1}_A$ integrable. $\operatorname{Int}(S\setminus A) = \emptyset$.

$$\mathbf{1}_{S\setminus A} = 0 \Rightarrow \int \mathbf{1}_A = \int \mathbf{1}_S \Rightarrow v(A) = v(S)$$

3.1 Improper Integrals

Problem 3.1.1. What if S and/or f not bounded?

Example. Newtonian potential $\int_{\mathbb{R}^3\setminus\{\vec{0}\}} \frac{f(\vec{x})}{\|\vec{x}\|}$.

Munkres starts to focus on $f^{\text{cont}} : A^{\text{open}} \to \mathbb{R}$ (See Munkres 15.8 for discontinuous f). For now, also assume $f \geq 0$. Define extended Riemman integral

$$\text{ext} \int_A f := \sup \left\{ \int_E f \mid E^{\text{compact, rect}} \subset A \right\}$$

Lemma 3.1.1. $B^{\text{open}} \subset A^{\text{open}}$ then $\text{ext} \int_B f \leq \text{ext} \int_A f$.

Answer. Assuming $f \geq 0$, we have the following

(i) old $\int_A f$ may not exist, A may not be rectifiable.

(ii) $E^{\text{compact, rect}} \subset A$ then

$$\int_E f = \underline{\int}_E f \leq \underline{\int}_A f = \underline{\int}_Q f$$

so

$$\text{ext} \int_A f \leq \underline{\int}_A f$$

(iii) P partition of $Q^{\text{box}} \supset A$ then

$$L(f_A, P) \leq \int_{\substack{\text{union of } P\text{-boxes} \\ \text{contained in A}}} f \leq \text{ext} \int_A f$$

Thus, $\underline{\int}_A = \text{ext} \int_A f$.

⊛

3.1.1 How to compute?

Suppose we have $E_1 \subset E_2 \subset \cdots \in A$ such that E_i compact rectifiable with $\bigcup_{i\in\mathbb{N}} \operatorname{Int} E_i = A$.

Example. $E_j = [-j, 0] \cup [\frac{1}{j}, j]$. Then $\bigcup E_j = \mathbb{R}$, $\operatorname{Int} E_j = \mathbb{R} \setminus \{0\}$.

Claim. $\text{ext} \int_A f = \lim_{j\to\infty} \int_{E_j} f$.

Proof. each $\int_{E_j} f \leq \text{ext} \int_A f$, so $\lim \int_{E_j} f \leq \text{ext} \int_A f$. For any $E \subset A$ compact, rectifiable, $E \subset E_j$ for some j. Therefore, $\int_E f \leq \int_{E_j} f \leq \lim \int_{E_j} f$, so $\text{ext} \int_A f \leq \lim \int_{E_j} f$. ■

Proposition 3.1.1 (Munkres (Lemma 15.1))**.** Can always find such E_j.

Alternative Proof Sketch. Let E_j be the union of all closed hypercubes with

- Side length $\frac{1}{j}$
- Cube $\subset A$

- Vertices have coordinates in $\frac{\mathbb{Z}}{2^j} \cap [-j, j]$.

As an exercise, show that this construction works. ■

Example. Consider the following

$$\begin{aligned}
\operatorname{ext} \int_{\mathbb{R}} \frac{1}{1+x^2} &= \lim_{j\to\infty} \int_{-j\le x\le j} \frac{1}{1+x^2} \\
&= \lim_{j\to\infty} \arctan j \Big|_{x=-j}^{x=j} \\
&= \lim_{j\to\infty} 2\arctan j \\
&= \pi
\end{aligned}$$

Example. Consider the following

$$\begin{aligned}
\int_{\mathbb{R}^2} \frac{1}{1+x^2+y^2} &= \lim_{j\to\infty} \int_{-j\le y\le j} \int_{-j\le x\le j} \frac{1}{1+x^2+y^2} \\
&= \lim_{j\to\infty} \int_{-j\le y\le j} \frac{2\arctan \frac{j}{\sqrt{1+y^2}}}{\sqrt{1+y^2}} \\
&= \lim_{j\to\infty} \lim_{k\to\infty} \int_{\substack{-j\le y\le j \\ -k\le x\le k}} \frac{1}{1+x^2+y^2} \\
&= \lim_{j\to\infty} \lim_{k\to\infty} \int_{-j\le y\le j} \frac{2\arctan \frac{k}{\sqrt{1+y^2}}}{\sqrt{1+y^2}} \\
&= \lim_{j\to\infty} \int_{-j\le y\le j} \frac{\pi}{\sqrt{1+y^2}}
\end{aligned}$$

The last step follows from uniform convergence on a bounded interval. Setting $y = \sinh u$,

$$\begin{aligned}
\lim_{j\to\infty} \int_{-j\le y\le j} \frac{\pi}{\sqrt{1+y^2}} &= \lim_{j\to\infty} \int_{-j\le y\le j} \frac{\pi}{\cosh u} \cosh u \\
&= \lim_{j\to\infty} \pi u \Big|_{u=-\operatorname{arcsinh} j}^{u=\operatorname{arcsinh} j} \\
&= \infty
\end{aligned}$$

Lecture 24: TBD

27 Oct. 02:30

Problem 3.1.2. Can you rewrite

$$I = \lim_{j\to\infty} \lim_{k\to\infty} \int_{\substack{-j\le y\le j \\ -k\le x\le k}} \frac{1}{1+x^2+y^2}$$

Claim. $\operatorname{ext} \int_A f = \lim_{j\to\infty} \int_{U_j} f$ for $U_j \subset$ open, such that $U_1 \subset U_2 \subset \cdots$ and $\bigcup U_j = A$.

Proof. Note that the RHS exists because it is a monotone increasing sequence. By monotinicity, LHS $\ge$ RHS. Now we will show that LHS $\le$ RHS. Note each

$$\operatorname{ext} \int_{U_j} f \le \operatorname{ext} \int_A f$$

Each compact rectifiable $E \subset A$ lies in some U_j, so $\int_E f \leq \operatorname{ext} \int_{U_j} f \leq \lim \operatorname{ext} \int_{U_j} f$. Now take sup over choice of E. ■

Remark. In the example, take $U_j = \mathbb{R} \times (-j, j)$.

We now want to get rid of the assumption that $f \geq 0$. For $x \in [-\infty, \infty]$, set $x_+ = \max\{x, 0\}$, $x_- = \max\{-x, 0\}$. Then

- $x_+, x_- \geq 0$,
- $x_+ \cdot x_- = 0$,
- $x = x_+ - x_-$,
- $|x| = x_+ + x_-$.

For $f : X \to [-\infty, \infty]$, let $f_+ : x \mapsto (f(x))_+$, $f_- : x \mapsto (f(x))_-$. Then

- $f = f_+ - f_-$,
- $|f| = f_+ + f_-$,
- $f_+ \cdot f_- = 0$,
- $|f| = f_+ + f_-$.

Now focus on $f \in C(A^{\text{osso } \mathbb{R}^n}, \mathbb{R})$.

Definition 3.1.1. f is *extended-integrable* or *integrable on A in the extended sense* if and only if $\operatorname{ext} \int_A f_+$ and $\operatorname{ext} \int_A f_- < \infty$. Equivalently, $\operatorname{ext} \int_A |f| = \operatorname{ext} \int_A f_+ + \operatorname{ext} \int_A f_- < \infty$. In which case, $\operatorname{ext} \int_A f = \operatorname{ext} \int_A f_+ - \operatorname{ext} \int_A f_-$.

Lecture 26: TBD

1 Nov. 02:30

Theorem 3.1.1. Given $A^{\text{osso } \mathbb{R}^n} \xrightarrow{g} B^{\text{osso } \mathbb{R}^n} \xrightarrow{f} \mathbb{R}$ for g diffeomorphism and f continuous. Then $\operatorname{ext} \int_B f = \operatorname{ext} \int_A (f \circ g) |\det Dg|$

On Monday, we needed the quantity $\int_{B^2(1)} (1 - \|x\|^2)^{\frac{n}{2}-1}$. Note that

$$
\begin{aligned}
\int_{B^2(1)} (1 - \|x\|^2)^{\frac{n}{2}-1} &= \int_{B^2(1) \setminus ((-1,0] \times \{0\})} (1 - \|x\|^2)^{\frac{n}{2}-1} + \int_{(-1,0] \times \{0\}} (1 - \|x\|^2)^{\frac{n}{2}-1} \\
&= \int_{\substack{0<r<1 \\ -\pi<\theta<\pi}} (1 - r^2)^{\frac{n}{2}-1} r \\
&= \frac{2\pi}{n}
\end{aligned}
$$

$B^n(r) = \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x}\| < r\}$, $\lambda_n = v(B^n(1)) = \frac{2\pi\lambda_{n-2}}{n}$ and $\lambda_{2n} = \frac{\pi^n}{n!}$. Now

$$
\lambda_n = \begin{cases} \frac{\pi^{n/2}}{(n/2)!}, & \text{if } n \text{ is even} \\ \frac{2^{\frac{n+1}{2}} \pi^{\frac{n-1}{2}}}{n(n-2)\cdots 3 \cdot 1}, & \text{if } n \text{ is odd} \end{cases}
$$

Let $\mu_n = v(\{\vec{x} \in \mathbb{R}^n \mid |\vec{x}| < 1\}) = 2^n$, $K_n = v\left(\{\vec{X} \in \mathbb{R}^n \mid |x_1| + \cdots + |x_n| < 1\}\right)$, so that $K_n \leq \lambda_n \leq \mu_n$. Check $\frac{\lambda_n}{\mu_n} \to 0$ and $\frac{K_n}{\lambda_n} \to 0$.

3.2 Proof of Change of Variables Theorem

3.2.1 First proof with temporary assumptions

Assume

- A, B bounded, rectifiable.
- g is a diffeomorphism from nbhd $\overline{A}$ to nbhd of $\overline{B}$.
- $f \in C(\overline{B}, \mathbb{R})$.

Proof of special case (1). Assume g is a coordinate transposition i.e.

$$g \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_k \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix}$$

Then $\det Dg = -1$ and $|\det Dg| = 1$. Since coordinate does not effect integration,

$$\text{ext} \int_B f = \text{ext} \int_A (f \circ g)$$

as desired. ∎

Proof of special case (2). Given $E^{\text{bdd open, rect}} \subset \mathbb{R}^{n-1}$. $\phi, \psi \in C(\overline{E}, \mathbb{R})$, $\phi < \psi$ on E. Define

$$B = \left\{ \vec{x} \;\middle|\; \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} \in E,\ \phi \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} < x_n < \psi \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} \right\}$$

and

$$g = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ \alpha(\vec{x}) \end{pmatrix}$$

Then $Dg = \begin{pmatrix} \vec{e}_1 \\ \vdots \\ \vec{e}_{n-1} \\ D\alpha \end{pmatrix}$ and $|\det Dg| = |D_n \alpha|$. Now define

$$A := \left\{ \vec{x} \;\middle|\; \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} \in E,\ \phi \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} < \alpha(\vec{x}) < \psi \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} \right\}$$

Then

$$\begin{aligned}
\int_B f &= \int_{\mathbf{x}^{(n-1)}\in E}\left(\int_{x_n\in(\phi(\mathbf{x}^{(n-1)}),\psi(\mathbf{x}^{(n-1)}))} f\right) \\
&= \int_{\mathbf{x}^{(n-1)}\in E}\left(\int_{\alpha(\vec{x})\in(\phi(\mathbf{x}^{(n-1)}),\psi(\mathbf{x}^{(n-1)}))} f\circ g|D_n\alpha|\right) \\
&= \int_A f\circ g|D_n\alpha| \\
&= \int_A f\circ g|\det Dg|
\end{aligned}$$

■

Proposition 3.2.1. Given $A \xrightarrow{g} B \xrightarrow{h} X \xrightarrow{f} \mathbb{R}$ for g, h diffemorphism and f continuous. If COVT holds for g, h, then COVT holds for $h\circ g$.

Proof. Note

$$\begin{aligned}
\int_C f &= \int_B (f\circ h)|\det Dh| \\
&= \int_A (f\circ h)\circ g\,(|\det Dh|\circ g)\,|\det Dg| \\
&= \int_A (f\circ h\circ g)|\det D(h\circ g)|
\end{aligned}$$

■

Strategy: Factor general diffeomorphism into composition of maps of type (1) and (2).

- Can do this!
- But only locally :(

Theorem 3.2.1. In some nbhd of $\vec{p}$, g can be factored into composition of diffeomorphisms of type (1) or (2).

Proposition 3.2.2. Any invertible affine map can be factored into a composition of affine maps of type (1) or (2).

Proof. Linear maps: Use "elementary matrix factorization". Translations: move one coordinate at a time. ■

Fun Fact!

- (1) is actually composition of 3 maps of type (3) = (1)(2)(1).

$$\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}$$

Proof. Use (1)-(2)-(1) to obtain

$$\begin{pmatrix} x_1 \\ \vdots \\ x_{j-1} \\ x_j \\ x_{j+1} \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_{j-1} \\ \eta(\vec{x}) \\ x_{j+1} \\ \vdots \\ x_n \end{pmatrix}$$

Remains to show for non-affine g.

Step 1: Pick $T_1, T_2 : \mathbb{R}^n \to \mathbb{R}^n$ invertible affine such that $g = T_1 \circ \widetilde{g} \circ T_2$ for $\widetilde{g} \coloneqq T_1^{-1} \circ g \circ T_2^{-1}$.

Then $T_2(\vec{p}) = 0$, $T_1(\vec{0}) = \vec{q}$, $DT_2 = Dg(\vec{p})$, $DT_1 = I$. Then $\widetilde{g}(\vec{0}) = \vec{0}$ and $D\widetilde{g}(\vec{0}) = I$, $D\widetilde{g}_k = \vec{e}_k$.

Step 2:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} \widetilde{g}_1(\vec{x}) \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} \widetilde{g}_1(\vec{x}) \\ \widetilde{g}_2(\vec{x}) \\ x_3 \\ \vdots \\ x_n \end{pmatrix},$$

By Inverse Function Theorem, both local diffeomorphism at $\vec{0}$. Consider

$$\begin{pmatrix} \widetilde{g}_1(\vec{x}) \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} \widetilde{g}_1(\vec{x}) \\ \widetilde{g}_2(\vec{x}) \\ x_3 \\ \vdots \\ x_n \end{pmatrix}$$

Then each of these is a type (3) map. Composing each of these gives you the mapping $\vec{x} \mapsto g(\vec{x})$ as desired.

$\blacksquare$

Lecture 27: TBD

Recall 6 Nov. 02:30

(1) Coordinate transposition

(2) $g_j(\vec{x}) = x_j$

(3) $g_j(\vec{x}) = x_j$ except for one point.

On Wednesday, we reduced to case $g(\vec{0}) = \vec{0}$.

Definition 3.2.1. $f : X^{\text{metric space}} \to V^{\text{vector space}}$. The *support* of f is the closure (in X) of $\{\vec{x} \mid f(\vec{x}) = \vec{0}\}$.

Remark. So $\vec{x} \in \operatorname{supp} f \Leftrightarrow \exists \epsilon > 0$ such that $f \equiv \vec{0} \in \mathcal{U}(\vec{x}, \epsilon)$.

Corollary 3.2.1. $\exists$ nbhd $\mathcal{U}$ of $\vec{a}$ such that COVT holds when $\operatorname{supp} f \subset \mathcal{U}$.

Theorem 3.2.2 (Paritition of Unity). Given $\Omega^{\text{open}} \subset \mathbb{R}^n$, $\Omega = \bigcup_{\alpha\in A} \mathcal{U}_\alpha^{\text{open}}$. Then $\exists \phi_1, \phi_2, \dots \in C^\infty(\Omega, [0, +\infty))$ such that

(i) Each $\operatorname{supp}\phi_j$ compact.

(ii) Each $\operatorname{supp}\phi_j \subset \mathcal{U}_{\alpha_j}$ for some α_j.

(iii) Each $\vec{x} \in \Omega$ has a nbhd meeting at only finitely many $\operatorname{supp}\phi_j$.

(iv) $\sum_{j=1}^\infty \phi_j(\vec{x}) = 1$ for all $\vec{x} \in \Omega$.

$\{\phi_j\}$ is a *partition of unity* on Ω dominated by $\{\mathcal{U}_\alpha\}$.

Proof. This will be proved Nov 20th. ■

Lemma 3.2.1. Given $f \in C(B^{\text{osso } \mathbb{R}^n}, \mathbb{R})$, ext $\int_B f$ exists and $\{\phi_j\}$ satisfies (i), (iii), (iv). Then ext $\int_B f = \sum_{j=1}^\infty \int_B \phi_j f$.

Proof. First cosider $f \geq 0$. Then for $E^{\text{compact, rect}} \subset B$, $\exists M$ such that $\phi_j \equiv 0$ on E for $j \geq M$. Therefore,

$$\int_E f = \sum_{j=1}^M \int_E \phi_j f \leq \sum_{j=1}^M \int_B \phi_j f$$

Take sup over all E, get

$$\text{ext} \int_B f \leq \sum_{j=1}^\infty \int_B \phi_j f$$

Also,

$$\begin{aligned} \sum_{j=1}^\infty \int_B \phi_j f &= \lim_{M\to\infty} \sum_{j=1}^\infty \int_B \phi_j f \\ &= \lim_{M\to\infty} \int_B \sum_{j=1}^M \phi_j f \\ &\leq \text{ext} \int_B f \end{aligned}$$

Finish by applying to f_+ and f_-. ■

Proof of COVT. For $\vec{y} \in B$ choose $\mathcal{U}_{\vec{y}}$ nbhd of $\vec{y}$ such that $\vec{g}$ factors on $\mathcal{U}_{\vec{y}}$. Choose partition of **1** $\{\phi_j\}$ dominated by $\{\mathcal{U}_{\vec{y}}\} \in B$. By Lemma,

$$\begin{aligned} \text{ext} \int_B f_+ = \sum \int_B \phi_j f_+ = \sum \int_A (\phi_j \circ g)(f_+ \circ g)|\det Dg| = \text{ext} \int_A (f_+ \circ g)|\det Dg| \\ \text{ext} \int_B f_- = \sum \int_B \phi_j f_- = \sum \int_A (\phi_j \circ g)(f_- \circ g)|\det Dg| = \text{ext} \int_A (f_- \circ g)|\det Dg| \end{aligned}$$

■

Can also consider:

- Length of curves in $\mathbb{R}^2, \mathbb{R}^3, \dots$
- Area of surfaces in $\mathbb{R}^3, \dots$
- k-volums of k-dimensional objects in $\mathbb{R}^n$.

Special role played by isometries of $\mathbb{R}^n$. $h : \mathbb{R}^n \to \mathbb{R}^n$ is an affine isometry if and only if $h(\vec{x}) = A\vec{x} + \vec{b}$ with $A^T A = I$.

Theorem 3.2.3. All isometries $h : \mathbb{R}^n \to \mathbb{R}^n$ are of this form.

Note that

$$v(h(\Omega)) = \int_{h(\Omega)} 1 = \int_{\Omega} |\det Dh| = \int_{\Omega} |\det A|$$

Since $A^T A = I$, we have that $\det A = \pm 1$. So $v(h(\Omega)) = v(\Omega)$.

Fact: Isometries are generated by

- Translations,
- Coordinate permutations,
- Rotations i.e. $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

Given $Q^{\text{box}} \subset \mathbb{R}^k \xrightarrow{T} \mathbb{R}^n$ with T affine injective. Then $T(Q)$ is a k-parallelopiped. We want a good definition of "k-vol of $T(Q)$". We have done this for $k = 1, n$. However, we do not have a definition in general.

Lecture 28: TBD

Want definition of v_k that satisfies 8 Nov. 02:30

1. For $A = \begin{pmatrix} M \\ 0 \end{pmatrix}$ For $M \in \mathbb{R}^{k \times k}$. Then $v_k(T(Q)) = |\det M| v(Q)$.
2. If $h : \mathbb{R}^n \to \mathbb{R}^n$ is an affine isometry then $v_k(h(T(Q))) = v_k(T(Q))$.

Claim. There is one unqiue map $v_k : \{k\text{-parallelopipeds}\} \to (0, +\infty)$ that satisfies (1) and (2).

Theorem 3.2.4. Can pick $B \in O(n) := \{\text{orthogonal matrices in } \mathrm{Mat}(n, n, \mathbb{R})\}$ such that $BA = \begin{pmatrix} M \\ 0 \end{pmatrix}$.

Proof Outline. First k rows are orthonormal bbasis for column space of A. Then extend to orthonormal basis for $\mathbb{R}^n$ for remaining $k - n$ rows. ■

Note.

$$(BA)^T = \begin{pmatrix} M^T & 0 \end{pmatrix}$$
$$(BA)^T \cdot (BA) = A^T B^T BA = A^T A = M^T M$$

If we assume (1), (2) then we now have

$$\begin{aligned} v_K(h(T(Q))) &= |\det M| v_k(Q) \\ &= \sqrt{|\det M^T M|} v_k(Q) \\ &= \sqrt{|\det A^T A|} v_k(Q) \end{aligned}$$

This leads us to the definition $v_k(T(Q)) = \sqrt{|\det A^T A|} v_k(Q)$.

Claim. $\ker A^T A = \ker A$

Proof. If $x \in \ker A$ then $x \in \ker A^T A$ trivially. For the reverse dircetion, if $A^T Ax = 0$ then

$$\|Ax\|^2 = x^T A^T Ax = 0$$

Since $\|Ax\| \geq 0$ we have $Ax = 0$. ■

Corollary 3.2.2. $\operatorname{rank} A^T A = \operatorname{rank} A$

Claim. $A^T A \geq 0$

Corollary 3.2.3. All eigenvalues of $A^T A$ are non-negative.

Corollary 3.2.4. $\det A^T A \geq 0$, strict if and only if $\operatorname{rank} A = k$.

Lecture 29: TBD

10 Nov. 02:30

Corollary 3.2.5. $T : [0, 1] \to \mathbb{R}^n$ defined by $t \mapsto (\vec{b} - \vec{a}) + \vec{a}$. Then

$$(1 - v(T([0, 1]))^2 = \sum(b_j - a_j)^2$$

Remark. If A "not special" then $\det A^T A = O(k^2(n + k))$ using Gaussian elimination. Using Pythagoras, it takes $\binom{n}{k}k^3$.

3.2.2 Non-affine

We will look at 3 k-dimensional objects.

- parametrized k-manifold.
- k-manifold in $\mathbb{R}^n$.
- "abstract" k-manifolds.

Parametrized k-manifold of class C^r $(r \geq 1)$:

- What it is - $\alpha \in C^r(\mathcal{U}^{\text{osso } \mathbb{R}^k}, \mathbb{R}^n)$.
- How we think about it - Focus on $y = \alpha(\mathcal{U})$ equipped with "coords" using α.
- Notation - Y_α

Define k-vol of Y_α by (see 22.4)

$$v_k(Y_\alpha) = \text{ext} \int_{\mathcal{U}} V(D_\alpha)$$

For $k = n$,

$$V(D_\alpha) = |\det D_\alpha|$$

Lecture 31: TBD

10 Nov. 02:30

Appendix

Appendix A

Additional Proofs