MATH 20800 (Honors Analysis in \mathbb{R}^n II) Notes

Steven Labalme

January 18, 2022

Contents

	The Algebra and Topology of \mathbb{R}^n 1.1 Notes	1 1
_	Differentiation 2.1 Notes	2
R	eferences	9

Chapter 1

The Algebra and Topology of \mathbb{R}^n

1.1 Notes

1/10:

- Syllabus.
 - In his mind, homework is the main setting where learning takes place.
- We're going to be studying analysis, or calculus, on manifolds this quarter.
- Manifold: A "space" that looks like Euclidean space \mathbb{R}^n locally.
 - The surfaces of a sphere and torus are common examples of 2-dimensional manifolds.
 - With regard to the above definition, think about how people in ancient times didn't think the Earth was a sphere because it looked like a plane locally.
- This class will look much like a calculus course, in that we first talk about limits, then differentiation, then integration, and culminating in the fundamental theory of calculus.
- Last quarter, we primarily developed linear algebra and basic topology on metric spaces.
 - Chapter 1 of Munkres (1991) is a review of what's needed from last quarter.
 - This is all basically continuity.
- Thus, we can start right up with differentiation.

Chapter 2

Differentiation

2.1 Notes

1/10:

• Since manifolds look like Euclidean spaces locally, we basically only need to study differentiation on Euclidean spaces.

• Set up: Let $U \subset \mathbb{R}^n$ be open, and $f: U \to \mathbb{R}^n$ be a function.

• Idea: The derivative of f at some point $\mathbf{a} \in U$ is "the best linear approximation" to f at \mathbf{a} .

• Differentiable (function f at \mathbf{a}): A function f for which there exists a linear transformation A: $\mathbb{R}^n \to \mathbb{R}^m$ such that

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

• Total derivative (of f at a): The linear transformation A corresponding to a differentiable function f. Denoted by Df(a).

• Questions to ask:

1. When does the total derivative exist?

2. When it does exist, can there be multiple?

3. When it exists and is unique, how do I calculate it?

• Proposition: If A, B are linear transformations that both satisfy the definition, then A = B.

- We have

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - A\mathbf{h}}{\|\mathbf{h}\|} = \mathbf{0} \qquad \qquad \lim_{\mathbf{h}\to\mathbf{0}} \frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - B\mathbf{h}}{\|\mathbf{h}\|} = \mathbf{0}$$

- It follows by subtracting the right equation above from the left one that

$$\lim_{\mathbf{h}\to\mathbf{0}}\frac{A\mathbf{h}-B\mathbf{h}}{\|\mathbf{h}\|}=\mathbf{0}$$

– Apply linearity: For $\mathbf{v} \in \mathbb{R}^n$ and $t \in \mathbb{R}$, t > 0, we have

$$\frac{A(t\mathbf{v}) - B(t\mathbf{v})}{t} = A\mathbf{v} - B\mathbf{v}$$

- Therefore, since $t\mathbf{v} \to 0$ as $t \to 0$, we have by the above that

$$\mathbf{0} = \lim_{t \to 0} \frac{A(t\mathbf{v}) - B(t\mathbf{v})}{\|t\mathbf{v}\|}$$

$$= \lim_{t \to 0} \frac{A\mathbf{v} - B\mathbf{v}}{\|\mathbf{v}\|}$$

$$\mathbf{0} \cdot \|\mathbf{v}\| = \lim_{t \to 0} (A\mathbf{v} - B\mathbf{v})$$

$$\mathbf{0} = A\mathbf{v} - B\mathbf{v}$$

$$B\mathbf{v} = A\mathbf{v}$$

- Example: Let $f: \mathbb{R}^n \to \mathbb{R}^m$ be linear, i.e., $f(\mathbf{v}) = A\mathbf{v}$ for some linear transformation A. Then for all $\mathbf{a} \in \mathbb{R}^n$, $Df(\mathbf{a}) = A$ is constant.
 - We have from the definition that

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

- Theorem: If f is differentiable at a, then f is continuous at a.
 - By definition, there exists a linear transformation A such that

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

- Additionally, we have that

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + A\mathbf{h} + \|\mathbf{h}\| \left(\frac{f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - A\mathbf{h}}{\|\mathbf{h}\|} \right)$$

- As $\mathbf{h} \to \mathbf{0}$, the right-hand side of the above equation goes to $f(\mathbf{a})$.
 - As a linear transformation, $A\mathbf{h} \to \mathbf{0}$ as $\mathbf{h} \to \mathbf{0}$.
 - Clearly $\|\mathbf{h}\| \to \mathbf{0}$ as $\mathbf{h} \to \mathbf{0}$.
 - And we have by definition that the last term goes to $\mathbf{0}$ as $\mathbf{h} \to \mathbf{0}$.
- Therefore, f is continuous at \mathbf{a} .
- Observation: A function $f: U \to \mathbb{R}^m$ is given by an m-tuple of functions $f_1: U \to \mathbb{R}$ known as components. $f = (f_1, \dots, f_m)$.
- Proposition: f is differentiable at $\mathbf{a} \in U$ iff each component function f_i is differentiable at \mathbf{a} . In this case,

$$Df(\mathbf{a}) = (Df_1(\mathbf{a}), \dots, Df_m(\mathbf{a}))$$

- We know that

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - A\mathbf{h}}{\|\mathbf{h}\|} \in \mathbb{R}^m$$

- Thus, the limit is zero iff the limit of each component is zero.
- We have that the i^{th} component of the vector on the left below is equal to the number on the right; we call the common value $L_i(\mathbf{h})$.

$$\left(\frac{f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - A\mathbf{h}}{\|\mathbf{h}\|}\right)_i = \frac{f_i(\mathbf{a} + \mathbf{h}) - f_i(\mathbf{a}) - (A\mathbf{h})_i}{\|\mathbf{h}\|}$$

- The upshot is that f is differentiable at \mathbf{a} iff $\lim_{\mathbf{h}\to\mathbf{0}} L_i(\mathbf{h}) = \mathbf{0}$ iff the linear transformation $\mathbf{h}\mapsto (A\mathbf{h})_i:\mathbb{R}^m\to\mathbb{R}$ is the total derivative of f_i .
- Now, each f_i is a function of n variables, i.e., $f_i(x_1,\ldots,x_n)$ where x_1,\ldots,x_n are coordinates on \mathbb{R}^n .
- Partial derivative (of f wrt. x_i at $\mathbf{a} \in U$): The following quantity. Denoted by $\partial f/\partial x_i$. Given by

$$\frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(\mathbf{a})}{h}$$

- The partial derivative is easy to calculate if you're good at calculating single-variable derivatives.
- Questions:

1/12:

- 1. If the partial derivatives all exist, does the total derivative also exist?
- 2. If partial derivatives exist, is f continuous?
- The answer is no to both it's too weak a condition.
 - Counter example: Consider $f: \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x,y) = \begin{cases} \frac{x^2 y}{x^4 + y^4} & (x,y) \neq \mathbf{0} \\ 0 & (x,y) = \mathbf{0} \end{cases}$$

- All partial derivatives exist at (0,0) but f is not continuous at (0,0).
- We'll consider this in the homework.
- Now we try taking derivatives in infinitely many directions, as opposed to just n many.
- Directional derivative (of f at \mathbf{a} in the direction of $\mathbf{v} \in \mathbb{R}^n$): The following quantity. Denoted by $D_{\mathbf{v}}f(\mathbf{a}), \partial f/\partial \mathbf{v}$. Given by

$$D_{\mathbf{v}}f(\mathbf{a}) = \frac{\partial f}{\partial \mathbf{v}} = \lim_{h \to 0} \frac{f(\mathbf{a} + h\mathbf{v}) - f(\mathbf{a})}{h}$$

- We always take $\|\mathbf{v}\| = 1$.
- The partial derivative is just a directional derivative along the standard basis vectors. Alternatively, the directional derivative is just a generalization of the partial derivatives.
- This still isn't a strong enough condition the above counterexample has all directional derivatives at (0,0) but still isn't continuous.
- Proposition: Suppose f is differentiable at $\mathbf{a} \in U$. Then all directional derivatives of f at \mathbf{a} exist and for all $\mathbf{v} \in \mathbb{R}^n$,

$$\frac{\partial f}{\partial \mathbf{v}} = Df(\mathbf{a})(\mathbf{v})$$

- The total derivative says that the derivative exists from all sequences of approach. We're just going to pick a particular vector direction of approach.
- Mathematically, by the definition of the total derivative,

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

• A particular consequence is that

$$\frac{\partial f}{\partial x_i} = Df(\mathbf{a})(e_i)$$

- But the total derivative, as a linear transformation, is completely defined by its behavior on the basis vectors.
- Thus, it is defined by the m-by-n matrix

$$Df(\mathbf{a}) = \left(\frac{\partial f_j}{\partial x_i}\right)_{\substack{1 \le j \le m \\ 1 \le i \le n}}$$

- Jacobian matrix (of f at a): The above matrix, representing the total derivative of f at a.
- Theorem: Suppose $f: U \to \mathbb{R}^m$ is a function on an open set $U \subset \mathbb{R}^n$. If all partial derivatives of f exist and are continuous on U, then f is differentiable on U.
 - Recall the mean value theorem (MVT): Suppose $g:[a,b]\to\mathbb{R}$ is a continuous function which is differentiable on (a,b). Then there exists $c\in(a,b)$ such that g'(c)=[g(b)-g(a)]/[b-a].
 - WLOG let m = 1 (if we prove this case, we can use the proposition relating f to its components to prove the general case).
 - Rewrite

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = f(a_1 + h_1, a_2 + h_2, a_3 + h_3, \dots, a_n + h_n) - f(a_1, a_2 + h_2, a_3 + h_3, \dots, a_n + h_n) + f(a_1, a_2 + h_2, a_3 + h_3, \dots, a_n + h_n) - f(a_1, a_2, a_3 + h_3, \dots, a_n + h_n) + \dots + f(a_1, \dots, a_{n-1}, a_n + h_n) - f(\mathbf{a})$$

where $\mathbf{a} = (a_1, ..., a_n)$ and $\mathbf{h} = (h_1, ..., h_n)$.

Apply the MVT to each term to get

$$f(a_1,\ldots,a_i+h_i,\ldots,a_n+h_n)-f(a_1,\ldots,a_i,\ldots,a_n+h_n)=h_i\frac{\partial f}{\partial x_i}(a_1,\ldots,c_i(\mathbf{h}),\ldots,a_n+h_n)$$

for some $c_i(\mathbf{h}) \in (a_i, a_i + h_i) \cup (a_i + h_i, a_i)$.

- Now let A be the Jacobian matrix of f at \mathbf{a} .
- WTS:

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

- We have

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a_1, \dots, c_i(\mathbf{h}), \dots, a_n + h_n)$$

- Let $\pi_i: \mathbb{R}^n \to \mathbb{R}^n$ be the linear map $(x_1, \dots, x_n) \mapsto (0, \dots, x_i, \dots, 0)$. Clearly, $\mathbf{x} = \sum_{i=1}^n \pi_i \mathbf{x}$.
- Thus, $A\mathbf{h} = \sum_{i=1}^{n} A\pi_i \mathbf{h}$ and $A\pi_i \mathbf{h} = \frac{\partial f}{\partial x_i}(\mathbf{a}) \cdot h_i$.
- Applying, we have

$$\frac{f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - A\mathbf{h}}{\|\mathbf{h}\|} = \sum_{i=1}^{n} \frac{1}{\|\mathbf{h}\|} \left(h_i \frac{\partial f}{\partial x_i} (a_1, \dots, a_{i-1}, c_i(\mathbf{h}), a_{i+1} + h_{i+1}, \dots, a_n + h_n) - \frac{\partial f}{\partial x_i} (\mathbf{a}) \cdot h_i \right)$$

$$= \sum_{i=1}^{n} \frac{h_i}{\|\mathbf{h}\|} \left(\frac{\partial f}{\partial x_i} (a_1, \dots, a_{i-1}, c_i(\mathbf{h}), a_{i+1} + h_{i+1}, \dots, a_n + h_n) - \frac{\partial f}{\partial x_i} (\mathbf{a}) \right)$$

- We know that $-1 \le h_i/\|\mathbf{h}\| \le 1$, so we need only show that the difference above goes to zero as $\mathbf{h} \to \mathbf{0}$. But we know this by the continuity of the partial derivatives.

- ullet Note that this theorem gives a sufficient condition but not a necessary condition for f to be differentiable.
- 1/14: Theorem (Chain Rule): Suppose $f: U \to \mathbb{R}^m$ and $g: V \to \mathbb{R}^p$ are functions defined on open sets $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ with $f(U) \subset V$. Suppose that f is differentiable at $\mathbf{a} \in U$ and g is differentiable at $\mathbf{b} = f(\mathbf{a}) \in V$. Then the composite function $g \circ f: U \to \mathbb{R}^p$ is differentiable at \mathbf{a} and $D(g \circ f)(\mathbf{a}) = Dg(\mathbf{b}) \circ Df(\mathbf{a}) : \mathbb{R}^n \to \mathbb{R}^p$.
 - Note that $f: U \to \mathbb{R}^m$ is differentiable at $\mathbf{a} \in U$ with derivative A iff $f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + A\mathbf{h} + \tilde{f}(\mathbf{h})$ such that

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{\tilde{f}(\mathbf{h})}{\|\mathbf{h}\|} = \mathbf{0}$$

where \tilde{f} is an error function.

- We're just rearranging terms here.
- If you like, \tilde{f} is the numerator from the definition of the total derivative.
- Let $A = Df(\mathbf{a}), B = Dg(\mathbf{b})$. Then

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + A\mathbf{h} + \tilde{f}(\mathbf{h})$$

so

$$\begin{split} (g \circ f)(\mathbf{a} + \mathbf{h}) &= g(f(\mathbf{a} + \mathbf{h})) \\ &= g(f(\mathbf{a})) + A\mathbf{h} + \tilde{f}(\mathbf{h}) \\ &= g(f(\mathbf{a})) + B(A\mathbf{h} + \tilde{f}(\mathbf{h})) + \tilde{g}(A\mathbf{h} + \tilde{f}(\mathbf{h})) \\ &= g(f(\mathbf{a})) + BA\mathbf{h} + B\tilde{f}(\mathbf{h}) + \tilde{g}(A\mathbf{h} + \tilde{f}(\mathbf{h})) \end{split}$$

- WTS: $\lim_{\mathbf{h}\to\mathbf{0}} [B\tilde{f}(\mathbf{h}) + \tilde{g}(A\mathbf{h} + \tilde{f}(\mathbf{h}))]/\|\mathbf{h}\| = \mathbf{0}.$
- For the first half of the fraction,

$$\frac{B\tilde{f}(\mathbf{h})}{\|\mathbf{h}\|} = B\left(\frac{\tilde{f}(\mathbf{h})}{\|\mathbf{h}\|}\right) \to \mathbf{0}$$

as $h \to 0$ since the argument goes to 0 as $h \to 0$ and B is a linear transformation (in particular, B(0) = 0).

- For the second half of the fraction,

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\tilde{g}(A\mathbf{h} + \tilde{f}(\mathbf{h}))}{\|\mathbf{h}\|} = \lim_{\mathbf{h}\to\mathbf{0}} \frac{\tilde{g}(A\mathbf{h} + \tilde{f}(\mathbf{h}))}{\|A\mathbf{h} + \tilde{f}(\mathbf{h})\|} \cdot \frac{\|A\mathbf{h} + \tilde{f}(\mathbf{h})\|}{\|\mathbf{h}\|}$$

- The left fraction on the right side of the equality goes to zero as $\mathbf{h} \to \mathbf{0}$ by the definition of \tilde{g} .
- The right fraction on the right side of the equality is bounded since

$$\frac{\left\|A\mathbf{h} + \tilde{f}(\mathbf{h})\right\|}{\|\mathbf{h}\|} \le \frac{\|A\mathbf{h}\|}{\|\mathbf{h}\|} + \frac{\left\|\tilde{f}(\mathbf{h})\right\|}{\|\mathbf{h}\|} \le \|A\| + \frac{\left\|\tilde{f}(\mathbf{h})\right\|}{\|\mathbf{h}\|}$$

where ||A|| is the operator norm and $||\tilde{f}(\mathbf{h})||/||\mathbf{h}|| \to 0$ as $\mathbf{h} \to \mathbf{0}$ by the definition of \tilde{f} .

- Thus, the second half of the fraction goes to zero as well.
- Theorem: Let $U \subset \mathbb{R}^m$ be an open subset.

1. Suppose $f, g: U \to \mathbb{R}^m$ are functions that are differentiable at $\mathbf{a} \in U$. Then f+g is also differentiable at $\mathbf{a} \in U$ and

$$D(f+g)(\mathbf{a}) = Df(\mathbf{a}) + Dg(\mathbf{a})$$

2. Suppose $f, g: U \to \mathbb{R}$ are both differentiable at $\mathbf{a} \in U$. Then $f \cdot g$ is also differentiable at \mathbf{a} , and

$$D(f \cdot g)(\mathbf{a}) = Df(\mathbf{a}) \cdot g(\mathbf{a}) + f(\mathbf{a}) \cdot Dg(\mathbf{a})$$

3. Suppose $f: U \to \mathbb{R}$ is differentiable at $\mathbf{a} \in U$ and $f(\mathbf{a}) \neq 0$. Then 1/f is differentiable at $\mathbf{a} \in U$ and

$$D(1/f)(\mathbf{a}) = -\frac{Df(\mathbf{a})}{f(\mathbf{a})^2}$$

- Proof of 1: Consider the functions $F: U \to \mathbb{R}^{2m}$ and $G: \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ defined by

$$F(\mathbf{x}) = (f(\mathbf{x}), g(\mathbf{x}))$$
 $G(\mathbf{y}, \mathbf{z}) = \mathbf{y} + \mathbf{z}$

so that

$$f + q = G \circ F$$

- lacksquare F is differentiable because its components are differentiable.
- G is differentiable because it's linear. This also implies that $DG(\mathbf{x}) = G$.
- Apply the chain rule to learn that $G \circ F$ is differentiable with derivative

$$D(f+g)(\mathbf{a}) = D(G \circ F)(\mathbf{a})$$

$$= DG(F(\mathbf{a})) \circ DF(\mathbf{a})$$

$$= G(DF(\mathbf{a}))$$

$$= G(Df(\mathbf{a}), Dg(\mathbf{a}))$$

$$= Df(\mathbf{a}) + Dg(\mathbf{a})$$

- Prove the others the same way.
- Theorem (Mean Value Theorem): Suppose $f: U \to \mathbb{R}$ is differentiable for all $\mathbf{a} \in U$ and that U contains the line segment joining $\mathbf{a}, \mathbf{a} + \mathbf{h} \in U$. Then there exists $t_0 \in (0,1)$ such that

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = Df(\mathbf{a} + t_0\mathbf{h})(\mathbf{h})$$

- Define $\phi(t) = f(\mathbf{a} + t\mathbf{h})$ for $t \in [0, 1]$.
- Apply the usual MVT to ϕ to learn that there exists $t_0 \in (0,1)$ such that $\phi(1) \phi(0) = \phi'(t_0)$.
- Then using the chain rule, $\phi'(t_0) = Df(\mathbf{a} + t_0\mathbf{h})(\mathbf{h})$.
- We now discuss higher order derivatives.
- Differentiable (f on U): A function f that is differentiable at every $\mathbf{a} \in U$.
- If f is differentiable on U, then the total derivative gives a map $Df: U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.
 - Note that $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is isomorphic to the set of all m-by-n matrices, and \mathbb{R}^{mn} .
- \bullet We can ask for Df to itself be differentiable. We define

$$D^2 f = D(Df)$$

if it exists and, more generally,

$$D^k f = D(D^{k-1} f)$$

• Class C^k (function): A function $f: U \to \mathbb{R}^m$ for which $Df, \dots, D^k f$ all exist and are continuous on U.

- Note that we technically need only require that $D^k f$ exist, as this implies the existence of $Df, \ldots, D^{k-1}f$.
- A function $f: U \to \mathbb{R}^m$ is of class C^k iff all partial derivatives $\partial f/\partial x_i: U \to \mathbb{R}^m$ exist and are of class C^{k-1} (this follows from the theorem relating partial derivatives and differentiability).
- Smooth (function): A function of class C^{∞} .

References

Munkres, J. R. (1991). Analysis on manifolds. Addison-Wesley.