

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 $f : U \rightarrow \mathbb{R}^m$ 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 $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that

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

- **Total derivative** (of f at \mathbf{a}): The linear transformation A corresponding to a differentiable function f . Denoted by $Df(\mathbf{a})$.
- 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

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

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

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

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

$$\begin{aligned}\mathbf{0} &= \lim_{t \rightarrow 0} \frac{A(t\mathbf{v}) - B(t\mathbf{v})}{\|t\mathbf{v}\|} \\ &= \lim_{t \rightarrow 0} \frac{A\mathbf{v} - B\mathbf{v}}{\|\mathbf{v}\|} \\ \mathbf{0} \cdot \|\mathbf{v}\| &= \lim_{t \rightarrow 0} (A\mathbf{v} - B\mathbf{v}) \\ \mathbf{0} &= A\mathbf{v} - B\mathbf{v} \\ B\mathbf{v} &= A\mathbf{v}\end{aligned}$$

- Example: Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ 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

$$\begin{aligned}\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - A\mathbf{h}}{\|\mathbf{h}\|} &= \lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{f(\mathbf{a}) + f(\mathbf{h}) - f(\mathbf{a}) - f(\mathbf{h})}{\|\mathbf{h}\|} \\ &= \lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\mathbf{0}}{\|\mathbf{h}\|} \\ &= \mathbf{0}\end{aligned}$$

- Theorem: If f is differentiable at \mathbf{a} , then f is continuous at \mathbf{a} .

- By definition, there exists a linear transformation A such that

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

- 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 $\mathbf{h} \rightarrow \mathbf{0}$, the right-hand side of the above equation goes to $f(\mathbf{a})$.

- As a linear transformation, $A\mathbf{h} \rightarrow \mathbf{0}$ as $\mathbf{h} \rightarrow \mathbf{0}$.
- Clearly $\|\mathbf{h}\| \rightarrow \mathbf{0}$ as $\mathbf{h} \rightarrow \mathbf{0}$.
- And we have by definition that the last term goes to $\mathbf{0}$ as $\mathbf{h} \rightarrow \mathbf{0}$.

- Therefore, f is continuous at \mathbf{a} .

- Observation: A function $f : U \rightarrow \mathbb{R}^m$ is given by an m -tuple of functions $f_1 : U \rightarrow \mathbb{R}$ 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

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

- 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 $\lim_{\mathbf{h} \rightarrow \mathbf{0}} L_i(\mathbf{h}) = \mathbf{0}$ iff the linear transformation $\mathbf{h} \mapsto (A\mathbf{h})_i : \mathbb{R}^m \rightarrow \mathbb{R}$ is the total derivative of f_i .

- Now, each f_i is a function of n variables, i.e., $f_i(x_1, \dots, x_n)$ where x_1, \dots, x_n are coordinates on \mathbb{R}^n .

1/12:

- **Partial derivative** (of f wrt. x_i at $\mathbf{a} \in U$): The following quantity. Denoted by $\partial f / \partial x_i$. Given by

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

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

- Questions:

1. 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 $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ 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

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

- 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,

$$\begin{aligned} \mathbf{0} &= \lim_{h \rightarrow 0} \frac{f(\mathbf{a} + h\mathbf{v}) - f(\mathbf{a}) - Df(\mathbf{a})(h\mathbf{v})}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(\mathbf{a} + h\mathbf{v}) - f(\mathbf{a})}{h} - Df(\mathbf{a})(\mathbf{v}) \\ Df(\mathbf{a})(\mathbf{v}) &= \frac{\partial f}{\partial \mathbf{v}} \end{aligned}$$

- 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

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

- **Jacobian matrix** (of f at \mathbf{a}): The above matrix, representing the total derivative of f at \mathbf{a} .
- Theorem: Suppose $f : U \rightarrow \mathbb{R}^m$ 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 $g : [a, b] \rightarrow \mathbb{R}$ 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

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

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

- Apply the MVT to each term to get

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

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:

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

- 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 $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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

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

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

- Note that this theorem gives a sufficient condition but not a necessary condition for f to be differentiable.