DIFFERENTIABILITY OF FUNCTIONS OF SEVERAL VARIABLES

In this lecture we will discuss the notion of differentiability of functions $f : \mathbb{R}^n \to \mathbb{R}^m$. Let us recall the definition of a differentiable function $f : \mathbb{R} \to \mathbb{R}$. We say that f is differentiable at $x \in \mathbb{R}$ if

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{1}$$

exists. We denote the value of the limit by $f'(x)$ and it is called the derivative of f at x.

1. Possible Definitions

1.1. **Direct Generalization.** Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a function, and let $X \in \mathbb{R}^n$. We want to see how we can define a quantity "$f'(X)$" which is consistent with the one variable case. The immediate thing which comes to our mind is the direct generalization of Equation (1). We have already learnt the concept of limits for functions from $\mathbb{R}^n$ to $\mathbb{R}^m$.

We say that f is differentiable at X if

$$\lim_{H \to 0} \frac{f(X+H) - f(X)}{H}$$

exists. Can you see that the above equation does not make any sense. Since $H \in \mathbb{R}^n$ is a vector, we cannot divide by H.

The immediate remedy is to take $||H||$ in place of H. So, the definition then says that f is differentiable at X if

$$\lim_{H \to 0} \frac{f(X+H) - f(X)}{||H||} \tag{2}$$

exists.

Let us consider the case when $m = n = 1$ and see whether this definition is equivalent to Equation (1). Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = |x|$. We already know that f is not differentiable at 0. The reason is that if we approach the point $x = 0$ from left-hand side, the limit in Equation (1) approaches -1, and if we approach $x = 0$ from right-hand side, the limit becomes 1. So, limit does not exist.

Let us check whether f is differentiable at 0 (according to our new definition, i.e., Equation (2)). Indeed,

$$\lim_{H \to 0} \frac{f(0+H) - f(0)}{||H||} = \lim_{H \to 0} \frac{|H|}{|H|} = 1.$$

So, f becomes differentiable at 0. This implies that this new definition cannot serve as a definition of differentiability of functions of several variables.

1.2. **Directional Derivatives.** We have seen in the last example that direction matters. So, to define differentiability in several variable we should be able to approach the point X from all possible directions. In other words, we want to measure the rate of change of a function at a point along a given direction. Consider a norm-one vector $U \in \mathbb{R}^n$. This define a direction in $\mathbb{R}^n$. We know that the parametric equation of a line passing through X in the direction of U is given by $X + tU$. As $t \to 0$, $X + tU \to X$. Consider the following definition.

Definition 1 (Directional Derivatives). *Let $f : \mathbb{R}^n \to \mathbb{R}^m$, $X \in \mathbb{R}^n$ and $U \in \mathbb{R}^n$ such that $||U|| = 1$. The* **directional derivative** *of f at X in the direction of U (or along U) is defined by*

$$D_U f(X) = \lim_{t\to 0} \frac{f(X+tU) - f(X)}{t} \tag{3}$$

provided limit exists.

Example. Let $f(x,y) = \frac{x^2y}{x^4+y^2}$ when $(x,y) \neq (0,0)$ and $f(0,0) = 0$. Let $U = (u_1, u_2) \in \mathbb{R}^2$ such that $||U|| = 1$, and $X = (0,0)$. Then

$$\frac{f(X+tU) - f(X)}{t} = \frac{f(tU) - f(0)}{t} = \frac{t^3u_1^2u_2}{t(t^4u_1^4 + t^2u_2^2)} = \frac{u_1^2u_2}{t^2u_1^4 + u_2^2}.$$

Thus,

$$D_{(u_1,u_2)}f(0,0) = \begin{cases} \frac{u_1^2}{u_2} & \text{if } u_2 \neq 0, \\ 0 & \text{if } u_2 = 0. \end{cases}$$

Remark 2. *(1) In case of functions of one-variable, we know that if a function is differentiable at a point, it must be continuous at that point.*

(2) In the above example, the directional derivative of f at $(0,0)$ exist in all directions, but f is not continuous at $(0,0)$ as we have seen in the previous lecture. You might wonder about the converse!

(3) The directional derivative of a function again cannot serve as a definition of differentiability, because we want our differentiable function to be at least continuous.

(4) Nevertheless, we will see later that directional derivatives is an important notion and it is intimately related to the actual definition of differentiability.

Example. Let $f(x,y) = |x| + |y|$ for $(x,y) \in \mathbb{R}$. Let $U = (u_1, u_2) \in \mathbb{R}^2$ be a unit vector and $X = (0,0)$. Then $|u_1| + |u_2| \neq 0$ and for any $t \in \mathbb{R}$ with $t \neq 0$, we have

$$\frac{f(X+tU) - f(X)}{t} = \frac{f(tU) - f(0)}{t} = \frac{|t|}{t}(|u_1| + |u_2|).$$

Hence, $D_{(u_1,u_2)}f(0,0)$ does not exist. Clearly, f is continuous at $(0,0)$, but none of the directional derivatives of f at $(0,0)$ exist.

1.3. **Partial Derivatives.** We have seen that the standard basis (or the standard unit vectors) $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, where 1 is at the i-th position and 0 elsewhere; $i = 1, 2, \ldots, n$, of $\mathbb{R}^n$ plays an important role in Linear Algebra. If we find the directional derivatives of a function f at a point X in the direction of the standard basis e_i, we call $D_{e_i}f(X)$ the i-th partial derivative of f at x. Let us consider the simple case of functions $f : \mathbb{R}^3 \to \mathbb{R}$.

Definition 3 (Partial Derivatives)**.** *Let $f : \mathbb{R}^3 \to \mathbb{R}$ and let $X_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$. The* **partial derivative** *of f with respect to x at X_0 is defined by*

$$\frac{\partial f}{\partial x}(x_0, y_0, z_0) = \lim_{h\to 0} \frac{f(x_0 + h, y_0, z_0) - f(x_0, y_0, z_0)}{h} \tag{4}$$

provided limit exists. Similarly, the partial derivatives of f at X_0 with respect to y and z are respectively defined by

$$\frac{\partial f}{\partial y}(x_0, y_0, z_0) = \lim_{k\to 0} \frac{f(x_0, y_0 + k, z_0) - f(x_0, y_0, z_0)}{k}$$

and

$$\frac{\partial f}{\partial z}(x_0, y_0, z_0) = \lim_{t\to 0} \frac{f(x_0, y_0, z_0 + t) - f(x_0, y_0, z_0)}{k}$$

provided the limits exist. The partial derivatives of f with respect to x, y and z are sometimes denoted by $f_x(x_0, y_0, z_0)$, $f_y(x_0, y_0, z_0)$ and $f_z(x_0, y_0, z_0)$, respectively.

Examples.

(1) Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x, y) = \sqrt{|xy|}$. Let us find the partial derivatives w.r.t. x and y at $(0, 0)$. Most of the students claim that $\frac{\partial f}{\partial x}(0, 0)$ and $\frac{\partial f}{\partial y}(0, 0)$ do not exist, because the derivative of f w.r.t x treating y as a constant is $\frac{\partial f}{\partial x} = \frac{1}{2}\sqrt{|\frac{y}{x}|}$ which is not defined at $(0, 0)$. This is incorrect.

We claim that $\frac{\partial f}{\partial x}(0, 0) = 0 = \frac{\partial f}{\partial y}(0, 0)$. Indeed,

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

Similarly, you can show that $\frac{\partial f}{\partial y}(0, 0) = 0$. It is also clear that f is continuous at $(0, 0)$. Try to find $D_U f(0, 0)$.

(2) Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x, y)\frac{xy}{x^2+y^2}$ for $(x, y) \neq (0, 0)$ and $f(0, 0) = 0$. Then

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h\to 0} \frac{f(0 + h, 0) - f(0, 0)}{h} = 0 \text{ and } \frac{\partial f}{\partial x}(0, 0) = \lim_{k\to 0} \frac{f(0, 0 + k) - f(0, 0)}{k} = 0.$$

We have already seen in the previous lecture that $\lim_{(x,y)\to(0,0)} f(x, y)$ does not exists, hence f is not continuous at $(0, 0)$.

(3) Let $f(x, y) = |x| + |y|$ for $(x, y) \in \mathbb{R}$. We have encountered this example earlier. None of the directional derivatives of f at $(0,)$ exist. Thus, $\frac{\partial f}{\partial x}(0, 0)$ and $\frac{\partial f}{\partial y}(0, 0)$ do not exist, but f is continuous at $(0, 0)$.

At this point you might wonder whether there is any relation between existence of partial derivatives and continuity of a function. The answer is in the next proposition.

Let us first recall the Mean Value Theorem.

Theorem 4. *Let $a, b \in \mathbb{R}$ with $a < b$. If $\phi : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b), then there is $c \in (a, b)$ such that*

$$\phi(b) - \phi(a) = \phi'(c)(b - a).$$

Proposition 5. *For any $f : [a, b] \times [c, d] \to \mathbb{R}$, we have the following:*

(1) If f_x exists on $[a, b] \times [c, d]$, then for each fixed $y_0 \in [c, d]$, the function $g : [a, b] \to \mathbb{R}$ given by $g(x) = f(x, y_0)$ is continuous.

(2) If f_y exists on $[a, b] \times [c, d]$, then for each fixed $x_0 \in [a, b]$, the function $h : [c, d] \to \mathbb{R}$ given by $h(y) = f(x_0, y)$ is continuous.

(3) If both f_x and f_y exist, and if one of them is bounded on $[a, b] \times [c, d]$, then f is continuous on $[a, b] \times [c, d]$.

Proof. (1) For fixed $y_0 \in [c, d]$, g is differentiable at $x_0 \in [a, b]$ if and only if f_x exists at (x_0, y_0) and $g'(x_0) = f_x(x_0, y_0)$. Hence, g is continuous at x_0.

(2) Similar to (1).

(3) Assume that both f_x and f_y exist, and f_x is bounded on $[a, b] \times [c, d]$. Then there exists $M \in \mathbb{R}$ such that $|f_x(u, v)| \leq M$ for all $(u, v)[a, b] \times [c, d]$. Fix $(x_0, y_0) \in [a, b] \times [c, d]$. Given any $(x, y) \in [a, b] \times [c, d]$ with $x \neq x_0$, by the MVT, there exists u between x and x_0 such that $f(x, y) - f(x_0, y) = f_x(u, y)(x - x_0)$. This implies that $|f(x, y) - f(x_0, y)| \leq M|x - x_0|$. Consequently,

$$\begin{aligned} |f(x, y) - f(x_0, y_0)| &\leq |f(x, y) - f(x_0, y)| + |f(x_0, y) - f(x_0, y_0)| \\ &\leq M|x - x_0| + |f(x_0, y) - f(x_0, y_0)|. \end{aligned}$$

These inequalities are valid if $x = x_0$. Thus, in view of (2), $f(x, y) \to f(x_0, y_0)$ as $(x, y) \to (x_0, y_0)$. So, f is continuous at (x_0, y_0).

□

2. Differentiability

It is clear from the previous discussion that the notion of differentiability of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is not a trivial one, and it cannot be directly generalized. If we want to generalize any concept, it is often helpful to come back and look at the concept more carefully. In this case, the definition of a differentiable function $f : \mathbb{R} \to \mathbb{R}$.

If f is differentiable at $x \in \mathbb{R}$, then $f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$. Using the $\epsilon - \delta$ definition of limit, for any given $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\left|\frac{f(x+h)-f(x)}{h} - f'(x)\right| < \epsilon \text{ whenever } 0 < |h| < \delta.$$

In other words,

$$\left|\frac{f(x+h)-f(x)-f'(x)h}{h}\right| < \epsilon \text{ whenever } 0 < |h| < \delta.$$

This implies that

$$\lim_{h\to 0} \frac{|f(x+h)-f(x)-f'(x)h|}{|h|} = 0. \tag{5}$$

If you look at Equation (5) carefully, can you see a linear map from $\mathbb{R} \to \mathbb{R}$? We already know from our Linear Algebra course that any linear map $A : \mathbb{R} \to \mathbb{R}$ is given by $A(x) = \alpha x$, for some $\alpha \in \mathbb{R}$. The linear map in Equation(5) is given by $A(h) = f'(x)h$. So, if $f : \mathbb{R} \to \mathbb{R}$ is differentiable at x, then there exists a linear map $A : \mathbb{R} \to \mathbb{R}$ given by $A(h) = f'(x)h$, such that Equation (5) is true. The converse is also true. In fact, if $A : \mathbb{R} \to \mathbb{R}$ is a linear map given by $A(x) = \alpha x$, for some $\alpha \in \mathbb{R}$, and

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

then

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

or

$$\lim_{h\to 0} \left|\frac{f(x+h)-f(x)-\alpha h}{h}\right| = \lim_{h\to 0} \left|\frac{f(x+h)-f(x)}{h} - \alpha\right| = 0.$$

It follows that,

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

Hence, f is differentiable and $f'(x) = \alpha$.

Thus, we arrive at the following result.

Proposition 6. *Let $f : \mathbb{R} \to \mathbb{R}$ be a function and let $x \in \mathbb{R}$. Then f is differentiable at x if and only if then there exists a linear map $A : \mathbb{R} \to \mathbb{R}$ such that*

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

We will imitate the condition in the above proposition in defining the concept of derivative of a function $f : \mathbb{R}^n \to \mathbb{R}^m$.

Definition 7 (Differentiability)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a function and let $X \in \mathbb{R}^n$. We say that f is differentiable at X if there exists a linear map $A : \mathbb{R}^n \to \mathbb{R}^m$ such that*

$$\lim_{h\to 0} \frac{||f(X+H)-f(X)-A(H)||}{||H||} = 0. \tag{6}$$

We call A the derivative of f at X, and denote it either by $f'(X)$ or by $Df(X)$.

If f is differentiable at every $X \in \mathbb{R}^n$, we say that f is differentiable on $\mathbb{R}^n$. Can you give an example of such a function?

Proposition 8. *Any linear map $A : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable, and $DA(X) = A$ for all $X \in \mathbb{R}^n$.*

Proof. The proof is very simple. Observe that

$$\lim_{h\to 0} \frac{||A(X+H) - A(X) - A(H)||}{||H||} = \lim_{h\to 0} \frac{||A(X) + A(H) - A(X) - A(H)||}{||H||} = 0.$$

□

Exercise 9. *(1) Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a constant function. Find the $Df(X)$.*

(2) Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be defined as $f(X) = A(X) + X_0$, where $A : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map and $X_0 \in \mathbb{R}^n$ is a fixed vector. Find $Df(X)$.

Exercise 10. *Prove that if f is differentiable at X, then A defined by Equation* (6) *is unique.*

We recommend first readers to skip the next example as well as the proof of Proposition 11.

Example. Consider the vector space $M_n(\mathbb{C})$ over $\mathbb{C}$. We know that $M_n(\mathbb{R})$ is an inner product space with the inner product $\langle A, B\rangle = trace(A^*B)$. This is called the Frobenius inner product. The norm induced by this inner product is given by

$$||A|| = \Big(trace(A^*A)\Big)^{1/2} = \Big(\sum_{i,j} |a_{ij}|^2\Big)^{1/2}, \text{where } A = (a_{ij}).$$

This norm is called the Frobenius norm or the Schatten 2-norm or the Hilbert-Schmidt norm. We can show that this norm is submultiplicative, i.e.,

$$||AB|| \le ||A||||B|| \text{ for all } A, B \in M_n(\mathbb{C}).$$

Consider the the map $f : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ given by $f(X) = X^2$. Since we have a norm on $M_n(\mathbb{C})$, we can define limit, continuity and differentiability of f exactly in the same way as we did for functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. We wish to show that f is differentiable on $M_n(\mathbb{C})$ and compute its derivative. Let $H \in M_n(\mathbb{C})$. We have

$$\begin{aligned} f(X+H) - f(X) &= (X+H)^2 - X^2 \\ &= X^2 + XH + HX + H^2 - X^2 \\ &= XH + HX + H^2. \end{aligned}$$

Can you see a linear map in the last equality. Define $A : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ by $A(H) = XH + HX$. Then A is linear. So, it remains to show that Equation (6) is satisfied. Indeed,

$$\lim_{H\to 0} \frac{||f(X+H) - f(X) - A(H)||}{||H||} = \lim_{H\to 0} \frac{||H^2||}{||H||} \leq \lim_{H\to 0} \frac{||H||^2}{||H||} = \lim_{H\to 0} ||H|| = 0.$$

Therefore, f is differentiable and $Df(X)(H) = XH + HX$. What happens if $n = 1$.

The next proposition gives us what we are expecting from any differentiable function.

Proposition 11. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $X \in \mathbb{R}^n$, then it is continuous at X.*

Proof. Let $\epsilon(H) = \frac{f(X+H)-f(X)-Df(X)(H)}{||H||}$. This is called the error function. Since f is differentiable at X, $\epsilon(H) \to 0$ as $H \to 0$. Now,

$$\begin{aligned} ||f(X+H) - f(X)|| &= ||Df(X)(H) + ||H||\epsilon(H)|| \\ &\leq ||Df(X)||||H|| + ||H||||\epsilon(H)|| \qquad (\star) \\ &\to 0 \text{ as } H \to 0. \end{aligned}$$

Therefore, f is continuous at X. □

Remark 12. *In the inequality ($\star$), we have used something that needs explanation. Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. We define the norm of A as*

$$||A|| = \sup_{\substack{x\in\mathbb{R}^n \\ ||x||=1}} ||Ax||. \tag{7}$$

We can show that for any linear map $A : \mathbb{R}^n \to \mathbb{R}^m$, we have $||A|| < \infty$. This implies that for all $x \in \mathbb{R}^n$, $||Ax|| \leq ||A||||x||$. Such a map is called a bounded linear map.