DIFFERENTIABILITY, DIRECTIONAL DERIVATIVES AND GRADIENT

In this lecture we will show that if a function is differentiable, then all its directional derivatives exist and they can be computed using the derivative of f.

1. Relation between Differentiability and Directional Derivatives

The next theorem tells us how we can recover the directional derivative from the derivative of a function.

Theorem 1. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $X \in \mathbb{R}^n$. Then $D_U f(X)$ exists for all $U \in \mathbb{R}^n$, $||U|| = 1$ and we have*

$$D_U f(X) = Df(X)(U). \tag{1}$$

Proof. Let $U \in \mathbb{R}^n$, $||U|| = 1$. Suppose $Df(X)$ exists. Then

$$\begin{aligned}
& f(X + tU) - f(X) - Df(X)(tU) = ||tU||\epsilon(tU) \\
\implies & f(X + tU) - f(X) - tDf(X)(U) = |t|\epsilon(tU), \ (\because Df(X) \text{ is linear}, ||U|| = 1) \\
\implies & \frac{f(X + tU) - f(X) - tDf(X)(U)}{t} = \frac{|t|}{t}\epsilon(tU), \\
\implies & \left|\left|\frac{f(X + tU) - f(X)}{t} - Df(X)(U)\right|\right| = \left|\left|\frac{|t|}{t}\epsilon(tU)\right|\right| = \epsilon(tU)|| \to 0 \text{ as } t \to 0, \\
& (\because t \to 0 \implies ||tU|| = |t| \to 0).
\end{aligned}$$

Therefore,

$$D_U f(X) = \lim_{t \to 0} \frac{f(X + tU) - f(X)}{t} = Df(X)(U).$$

□

2. Jacobian Matrix

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable function at $X \in \mathbb{R}^n$. Then the derivative $Df(X) : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map. We know that any such linear map with respect to the standard (or canonical) basis is represented by an $m \times n$ matrix, say M, and $Df(X)(v) = Mv$. Here, we are considering elements of $\mathbb{R}^n$ as column vectors. Let us try to identify the matrix M.

Let $\{e_1, \ldots, e_n\}$ and $\{v_1, \ldots, v_m\}$ be the standard (or canonical) basis of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Then the j-th column of M is given by

$$Me_j = Df(X)(e_j) = D_{e_j} f(X) = \frac{\partial f}{\partial x_j}(X). \tag{2}$$

Now, for any $X \in \mathbb{R}^n$, $f(X) \in \mathbb{R}^m$, so it has m components. Thus, we can write $f(X) = (f_1(X), f_2(X), \ldots, f_m(X)$, where $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$ are called the coordinate functions of f. Sometime we write $f = (f_1, f_2, \ldots, f_m)$. One can show that convergence of function from $\mathbb{R}^n$ to $\mathbb{R}^m$ reduces to coordinate-wise convergence. So, if we consider f as a column vector, it can be shown that

$$\frac{\partial f}{\partial x_j}(X) = \begin{pmatrix} \frac{\partial f_1}{\partial x_j}(X) \\ \frac{\partial f_2}{\partial x_j}(X) \\ \vdots \\ \frac{\partial f_m}{\partial x_j}(X) \end{pmatrix}$$

Therefore, the matrix of $Df(X)$ becomes

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X) & \frac{\partial f_1}{\partial x_2}(X) & \cdots & \frac{\partial f_1}{\partial x_n}(X) \\ \frac{\partial f_2}{\partial x_1}(X) & \frac{\partial f_2}{\partial x_2}(X) & \cdots & \frac{\partial f_2}{\partial x_n}(X) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(X) & \frac{\partial f_m}{\partial x_2}(X) & \cdots & \frac{\partial f_m}{\partial x_n}(X) \end{pmatrix}.$$

This matrix is denoted by $Jf(X)$, and is called the Jacobian matrix of f at X.

Gradient. In the case when $f : \mathbb{R}^n \to \mathbb{R}$, the **Jacobian matrix of f at X** becomes

$$\left(\frac{\partial f}{\partial x_1}(X), \frac{\partial f}{\partial x_2}(X), \cdots, \frac{\partial f}{\partial x_n}(X) \right).$$

This vector is called the **gradient of f at X** and is denoted by $\nabla f(X)$.

3. Functions of Two Variables

In this section we will focus on functions $f : \mathbb{R}^2 \to \mathbb{R}$. If f is differentiable at $X = (x_0, y_0)$, then the derivative $Df(X)$ is given by $\nabla f(X) = \left(\frac{\partial f}{\partial x}(X), \frac{\partial f}{\partial y}(X)\right)$. Thus, in order to prove that the function f is differentiable at (x_0, y_0), we have to show that the error function

$$\epsilon(h, k) = \frac{f(x_0 + h, y_0 + k) - f(x_0, y_0) - \nabla f(x_0, y_0)(h, k)}{||(h, k)||} \to 0,$$

as $(h, k) \to (0, 0)$.

Examples.

(1) Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x, y) = \sqrt{|xy|}$. We have seen that $\nabla f(0, 0) = (0, 0)$. Now,

$$\epsilon(h, k) = \frac{f(h, k) - f(0, 0) - \nabla f(0, 0)(h, k)}{||(h, k)||} = \frac{\sqrt{|hk|}}{\sqrt{h^2 + k^2}} \nrightarrow 0 \text{ as } (h, k) \to (0, 0).$$

Hence, f is not differentiable at $(0, 0)$.

We can show this using Theorem 1 also. Given a unit vector (u_1, u_2) in $\mathbb{R}^2$ and any $t \in \mathbb{R}$ with $t \neq 0$, we have

$$\frac{f(0+tu_1, 0+tu_2) - f(0,0)}{t} = \frac{|t|\sqrt{|u_1u_2|}}{t}.$$

It follows that the directional derivative $D_{(u_1,u_2)}f(0,0)$ does not exist whenever u_1 and u_2 are nonzero. Hence, by Theorem 1 we conclude that f is not differentiable at $(0,0)$.

(2) Let

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

Verify that $\nabla f(0,0) = (0,0)$. Thus,

$$\begin{aligned} |\epsilon(h,k)| &= \left|\frac{f(h,k) - f(0,0) - \nabla f(0,0)(h,k)}{||(h,k)||}\right| = \left|\frac{hk(h^2-k^2)}{(h^2+k^2)\sqrt{h^2+k^2}}\right| \\ &\leq \left|\frac{hk}{\sqrt{h^2+k^2}}\right| \leq \frac{|h^2+k^2|}{2\sqrt{h^2+k^2}} \\ &= \frac{\sqrt{h^2+k^2}}{2} \to 0 \text{ as } (h,k) \to (0,0). \end{aligned}$$

Hence, f is differentiable at $(0,0)$ and $Df(0,0) = (0,0)$.

Problem 2. *Define the function* $f : \mathbb{R}^2 \to \mathbb{R}$ *by*

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

Answer the following questions.

(1) *Discuss the continuity of* f *at* $(0,0)$.
(2) *Evaluate* $f_y(x,0)$ *for* $x \neq 0$.
(3) *Is* f_y *continuous at* $(0,0)$.
(4) *Find all directional derivative of* f *at* $(0,0)$.
(5) *Discuss the differentiability of* f *at* $(0,0)$.

Solution. (1) $|f(x,y) - f(0,0)| = |\frac{x^2(x-y)}{x^2+y^2}| \leq |x-y| \longrightarrow 0$ as $(x,y) \longrightarrow 0$.
Thus, f is continuous at $(0,0)$.

(2)

$$\begin{aligned} f_y(x,0) &= \lim_{h\to 0} \frac{f(x,h) - f(x,0)}{h} \\ &= \lim_{h\to 0} \frac{\frac{x^2(x-h)}{x^2+h^2} - x}{h} = -1. \end{aligned}$$

(3) $f_y(0,0) = \lim_{h\to 0} \dfrac{f(0,h) - f(0,0)}{h} = 0.$
Since $f_y(x,0) \not\to f_y(0,0)$ as $x \to 0$, f_y is not continuous at $(0,0)$.

(4) The directional derivative of f at $(0,0)$ in the direction of (u_1, u_2) is

$$\begin{aligned} D_{(u_1,u_2)} f(0,0) &= \lim_{t\to 0} \frac{f((0,0) + t(u_1,u_2)) - f(0,0)}{t} \\ &= \lim_{t\to 0} \frac{t^2 u_1^2 (t(u_1 - u_2))}{t(t^2 u_1^2 + t^2 u_2^2)} \\ &= \frac{u_1^2(u_1 - u_2)}{u_1^2 + u_2^2}. \end{aligned}$$

(5) $f_x(0,0) = \lim_{h\to 0} \dfrac{f(h,0) - f(0,0)}{h} = 1$. Thus, $\nabla f(0,0) = (1,0)$.
Let $H = (h,k) \in \mathbb{R}^2$. Then,

$$\begin{aligned} |\epsilon(H)| &= \left| \frac{f(h,k) - f(0,0) - h f_x(0,0) - k f_y(0,0)}{||H||} \right| \\ &= \left| \frac{\frac{h^2(h-k)}{h^2+k^2} - h}{\sqrt{h^2 + k^2}} \right| \\ &= \left| \frac{hk(h+k)}{(h^2+k^2)^{3/2}} \right|. \end{aligned}$$

Take $h = k$. Then $\epsilon(h,h) = \frac{1}{\sqrt{2}} \not\to 0$ as $h \to 0$. Therefore, f is not differentiable at $(0,0)$.

□

In the above example, we have seen that f is continuous at $(0,0)$ and all directional derivatives of f at $(0,0)$ exist, but f is not differentiable at $(0,0)$. So, at this point it is natural to ask under what assumptions on the directional (or partial) derivatives the function becomes differentiable. The following criterion answer this question.

Theorem 3. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be such that both f_x and f_y exist in a neighborhood of (x_0, y_0), and one of them is continuous at (x_0, y_0), then f is differentiable at (x_0, y_0).*