Lecture 28 : Directional Derivatives, Gradient, Tangent Plane

The partial derivative with respect to x at a point in $\mathbb{R}^3$ measures the rate of change of the function along the X-axis or say along the direction $(1,0,0)$. We will now see that this notion can be generalized to any direction in $\mathbb{R}^3$.

Directional Derivative : Let $f:\mathbb{R}^3\to\mathbb{R}$, $X_0\in\mathbb{R}^3$ and $U\in\mathbb{R}^3$ such that $\parallel U\parallel=1$. The directional derivative of f in the direction U at $X_0=(x_0,y_0,z_0)$ is defined by

$$D_{X_0}f(U)=\lim_{t\to 0}\frac{f(X_0+tU)-f(X_0)}{t}$$

provided the limit exists.

It is clear that $D_{X_0}f(e_1)=f_x(X_0)$, $D_{X_0}f(e_2)=f_y(X_0)$ and $D_{X_0}f(e_3)=f_z(X_0)$.

The proof of the following theorem is similar to the proof of Theorem 26.2.

Theorem 28.1: *If f is differentiable at X_0, then $D_{X_0}f(U)$ exists for all $U\in\mathbb{R}^3$, $\parallel U\parallel=1$. Moreover, $D_{X_0}f(U)=f'(X_0)\cdot U=(f_x(X_0),f_y(X_0),f_z(X_0))\cdot U$.*

The previous theorem says that if a function is differentiable then all its directional derivatives exist and they can be easily computed from the derivative.

Examples :

(i) In this example we will see that a function is not differentiable at a point but the directional derivatives in all directions at that point exist.

Define $f:\mathbb{R}^2\to\mathbb{R}$ by $f(x,y)=\frac{x^2y}{x^4+y^2}$ when $(x,y)\neq(0,0)$ and $f(0,0)=0$.

This function is not continuous at $(0,0)$ and hence it is not differentiable at $(0,0)$.

We will show that the directional derivatives in all directions at $(0,0)$ exist. Let $U=(u_1,u_2)\in\mathbb{R}^3$, $\parallel U\parallel=1$ and $\mathbf{0}=(0,0)$. Then

$$\lim_{t\to 0}\frac{f(\mathbf{0}+tU)-f(\mathbf{0})}{t}=\lim_{t\to 0}\frac{t^3u_1^2u_2}{t(t^4u_1^4+t^2u_2^2)}=\lim_{t\to 0}\frac{u_1^2u_2}{t^2u_1^4+u_2^2}=0,\text{ if } u_2=0 \text{ and } \frac{u_1^2}{u_2},\text{ if } u_2\neq 0$$

Therefore, $D_{\mathbf{0}}f((u_1,0))=0$ and $D_{\mathbf{0}}f((u_1,u_2))=\frac{u_1^2}{u_2}$ when $u_2\neq 0$.

(ii) In this example we will see that the directional derivative at a point with respect to some vector may exist and with respect to some other vector may not exist.

Consider the function $f(x,y)=\frac{x}{y}$ if $y\neq 0$ and 0 if $y=0$. Let $U=(u_1,u_2)$ and $\parallel U\parallel=1$. It is clear that if $u_1=0$ or $u_2=0$, then $D_{\mathbf{0}}f(U)$ exists and is equal to 0. If $u_1u_2\neq 0$ then

$$\lim_{t\to 0}\frac{f(\mathbf{0}+tU)-f(\mathbf{0})}{t}=\lim_{t\to 0}\frac{u_1}{tu_2}$$

does not exist. So, only the partial derivatives of the function at $\mathbf{0}$ exist. Note that this function can not be differentiable at $\mathbf{0}$ (Why ?).

Problem 1: *Let $f(x,y)=\frac{y}{|y|}\sqrt{x^2+y^2}$ if $y\neq 0$ and $f(x,y)=0$ if $y=0$. Show that f is continuous at $(0,0)$, it has all directional derivatives at $(0,0)$ but it is not differentiable at $(0,0)$.*

Solution : Note that $\mid f(x,y)-f(0,0)\mid=\sqrt{x^2+y^2}$. Hence the function is continuous.

For $\| (u_1, u_2) \| = 1$, $\lim_{t\to 0} \frac{f(tu_1,tu_2)}{t} = 0$ if $u_2 = 0$ and $\frac{u_2}{|u_2|}$ if $u_2 \neq 0$. Therefore directional derivatives in all directions exist.

Note that $f_x(0,0) = 0$ and $f_y(0,0) = 1$. If f is differentiable at $(0,0)$ then $f'(0,0) = \alpha = (0,1)$. Note that

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

For example, $h = k$ gives $(\sqrt{2}-1)\frac{k}{|k|} \nrightarrow 0$ as $k \to 0$. Therefore the function is not differentiable at $(0,0)$. $\square$

The vector $(f_x(X_0), f_y(X_0), f_z(X_0))$ is called **gradient** of f at X_0 and is denoted by $\nabla f(X_0)$.

An Application : Let us see an application of Theorem 1. Suppose f is differentiable at X_0. Then $f'(X_0) = \nabla f(X_0)$ and $D_{X_0} f(U) = \nabla f(X_0) \cdot U = \| \nabla f(X_0) \| \cos\theta$ where $\theta \in [0,\pi]$ is the angle between the gradient and U. Suppose $\nabla f(X_0) \neq 0$. Then $D_{X_0} f(U)$ is maximum when $\theta = 0$ and minimum $\theta = \pi$. That is, f increases (respectively, decreases) most rapidly around X_0 in the direction $U = \frac{\nabla f(X_0)}{\|\nabla f(X_0)\|}$ (respectively, U = -$\frac{\nabla f(X_0)}{\|\nabla f(X_0)\|}$).

Example: Suppose the temperature of a metallic sheet is given as $f(x,y) = 20 - 4x^2 - y^2$. We will start from the point $(2,1)$ and find a path i.e., a plane curve, $r(t) = x(t)i + y(t)j$ which is a path of maximum increase in the temperature. Note that the direction of the path is $r'(t)$. This direction should coincide with that of the maximum increase of f. Therefore, $\alpha r'(t) = \nabla f$ for some α. This implies that $\alpha x'(t) = -8x$ and $\alpha y'(t) = -2y$. By chain rule we have $\frac{dy}{dx} = \frac{2y}{8x} = \frac{y}{4x}$. Since the curve passes through (2,1), we get $x = 2y^4$.

We will now see a geometric interpretation of the derivative i.e, gradient.

Tangent Plane: Suppose $f : \mathbb{R}^3 \to \mathbb{R}$ is differentiable and $c \in \mathbb{R}$. Consider the surface $S = \{(x,y,z) : f(x,y,z) = c\}$. This surface is called a level surface at the height c. (For example if $f(x,y,z) = x^2 + y^2 + z^2$ and $c = 1$, then S is the unit sphere.) Let $P = (x_0, y_0, z_0)$ be a point on S and $R(t) = (x(t), y(t), z(t))$ be a differentiable (i.e., smooth) curve lying on S. With these assumptions we prove the following result.

Theorem 28.2: *If T is the tangent vector to $R(t)$ at P then $\nabla f(P) \cdot T = 0$.*

Proof : Since $R(t)$ lies on S, $f(x(t), y(t), z(t)) = c$. Hence $\frac{df}{dt} = 0$. By chain rule,

$$\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} = 0 \quad i.e., \ \nabla f \cdot \frac{dR}{dt} = 0 \quad i.e., \ \nabla f \cdot T = 0 \text{ at } P. \qquad \square$$

From the previous theorem we conclude the following. Note that the gradient $\nabla f(P)$ is perpendicular to the tangent vector to every smooth curve $R(t)$ on S passing through P. That is, all these tangent vectors lie on a plane which is perpendicular to $\nabla f(P)$. That is, $\nabla f(P)$, when $\nabla f(P) \neq 0$, is the normal to the surface at P. Therefore, the plane through P with normal $\nabla f(P)$ defined by

$$f_x(P)(x - x_0) + f_y(P)(y - y_0) + f_z(P)(z - z_0) = 0$$

is called the tangent plane to the surface S at $P = (x_0, y_0, z_0)$.

Suppose the surface is given as a graph of $f(x,y)$, i.e., $S = \{(x, y, f(x,y)) : (x,y) \in D \subseteq \mathbb{R}^2\}$. Then it can be considered as a level surface $S = \{(x,y,z) : F(x,y,z) = 0\}$ where $F(x,y,z) = f(x,y) - z$. Let $X_0 = (x_0, y_0)$, $z_0 = f(x_0, y_0)$ and $P = (x_0, y_0, z_0)$. Then the equation of the tangent plane is $f_x(x_0,y_0)(x - x_0) + f_y(x_0,y_0)(y - y_0) - (z - z_0) = 0$ i.e.,

$$z = f(X_0) + f'(X_0)(X - X_0), \quad X = (x,y) \in \mathbb{R}^2.$$