Lectures 26-27: Functions of Several Variables
(Continuity, Differentiability, Increment Theorem and Chain Rule)

The rest of the course is devoted to calculus of several variables in which we study continuity, differentiability and integration of functions from $\mathbb{R}^n$ to $\mathbb{R}$, and their applications.

In calculus of single variable, we had seen that the concept of convergence of sequence played an important role, especially, in defining limit and continuity of a function, and deriving some properties of $\mathbb{R}$ and properties of continuous functions. This motivates us to start with the notion of convergence of a sequence in $\mathbb{R}^n$. For simplicity, we consider only $\mathbb{R}^2$ or $\mathbb{R}^3$. General case is entirely analogous.

Convergence of a sequence : Let $X_n = (x_{1,n}, x_{2,n}, x_{3,n}) \in \mathbb{R}^3$. We say that the sequence (X_n) is convergent if there exists $X_0 \in \mathbb{R}^3$ such that $\|X_n - X_0\| \to 0$ as $n \to \infty$. In this case we say that X_n converges to X_0 and we write $X_n \to X_0$.

Note that corresponding to a sequence (X_n), $X_n = (x_{1,n}, x_{2,n}, x_{3,n})$, there are three sequences $(x_{1,n})(x_{2,n})$ and $(x_{3,n})$ in $\mathbb{R}$, and vice-versa. Thus the properties of (X_n) can be completely understood in terms of the properties of the corresponding sequences $(x_{1,n})(x_{2,n})$ and $(x_{3,n})$ in $\mathbb{R}$. For example,

(i) $X_n \to X_0$ in $\mathbb{R}^3 \Leftrightarrow$ the coordinates $x_{i,n} \to x_{i,0}$ for every $i = 1, 2, 3$ in $\mathbb{R}$.

(ii) (X_n) is bounded (i.e., $\exists\ M$ such that $\|X_n\| \le M\ \forall\ n$) $\Leftrightarrow$ each sequence $(x_{i,n}), i = 1, 2, 3$, is bounded.

Using the previous idea, we can prove the following results.

Problem 1: *Every convergent sequence* $\mathbb{R}^3$ *is bounded.*

Problem 2 (Bolzano-Weierstrass Theorem): *Every bounded sequence in* $\mathbb{R}^3$ *has a convergent subsequence.*

In case of a sequence in $\mathbb{R}$, to define the notion of convergence or boundedness, we use $|\ \ |$ in place of $\|\ \ \|$, hence it is clear how we generalized the concept of convergence or boundedness of a sequence in $\mathbb{R}^1$ to $\mathbb{R}^3$. Moreover, it is also now clear how to define the concepts of limit and continuity of a function $f : \mathbb{R}^3 \to \mathbb{R}$ at some point $X_0 \in \mathbb{R}^3$.

Limit and Continuity : (i) We say that L is the limit of a function $f : \mathbb{R}^3 \to \mathbb{R}$ at $X_0 \in \mathbb{R}^3$ (and we write $\lim_{X \to X_0} f(X) = L$) if $f(X_n) \to L$ whenever a sequence (X_n) in $\mathbb{R}^3$, $X_n \neq X_0$, converges to X_0.

(ii) A function $f : \mathbb{R}^3 \to \mathbb{R}$ is continuous at $X_0 \in \mathbb{R}^3$ if $\lim_{X \to X_0} f(X) = f(X_0)$.

Examples 1: (i) Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$, where $f(x, y) = \frac{\sin^2(x-y)}{|x|+|y|}$ when $(x, y) \neq (0, 0)$ and $f(0, 0) = 0$. We will show that this function is continuous at (0,0). Note that

$$|\ f(x, y) - f(0, 0)\ | \le \frac{|\ x - y\ |^2}{|\ x\ | + |\ y\ |}\ \le |\ x\ | + |\ y\ |\ \ (\text{or } |\ x - y\ |\ )$$

Therefore, whenever a sequence $(x_n, y_n) \to (0, 0)$, i.e, $x_n \to 0$ and $y_n \to 0$, we have $f(x_n, y_n) \to f(0, 0)$. Hence f is continuous at $(0, 0)$. In fact, this function is continuous on the entire $\mathbb{R}^2$.

(ii) Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$, where $f(x, y) = \frac{xy}{\sqrt{x^2+y^2}}$ when $(x, y) \neq (0, 0)$ and $f(0, 0) = 0$. This function is continuous at $(0, 0)$, because, $|\frac{xy}{\sqrt{x^2+y^2}}| \le \frac{|x^2+y^2|}{\sqrt{x^2+y^2}} = \sqrt{x^2 + y^2} \to 0$, as $(x, y) \to 0$.

(iii) Let $f(x,y) = \frac{2xy}{x^2+y^2}, (x,y) \neq (0,0)$. We will show that this function does not have a limit at $(0,0)$. Note that $f(x,mx) \to \frac{2m}{1+m^2}$ as $x \to 0$ for any m. This shows that the function does not have a limit at $(0,0)$.

(iv) Let $f(x,y) = \frac{x^2y}{x^4+y^2}$ when $(x,y) \neq (0,0)$ and $f(0,0) = 0$. Note that $f(x,mx) \to 0$ as $x \to 0$. But the function is not continuous at $(0,0)$ because $f(x,x^2) \to \frac{1}{2}$ as $x \to 0$. Similarly we can show that the function $f(x,y)$ defined by $f(x,y) = \frac{x^4-y^2}{x^4+y^2}$ when $(x,y) \neq (0,0)$ and $f(0,0) = 0$ is not continuous at $(0,0)$ by taking $y = mx^2$ and allowing $x \to 0$.

Partial derivatives : The partial derivative of f with respect to the first variable at $X_0 = (x_0, y_0, z_0)$ is defined by

$$\frac{\partial f}{\partial x} |_{X_0} = \lim_{h \to 0} \frac{f(x_0+h, y_0, z_0) - f(x_0, y_0, z_0)}{h}$$

provided the limit exists. Similarly we define $\frac{\partial f}{\partial y} |_{X_0}$ and $\frac{\partial f}{\partial z} |_{X_0}$.

Example 2: The function f defined by $f(x,y) = \frac{2xy}{x^2+y^2}$ *at* $(x,y) \neq (0,0)$ and $f(0,0) = 0$ is not continuous at $(0,0)$, however, the partial derivatives exist at $(0,0)$.

Problem 3: *Let $f(x,y)$ be defined in $S = \{(x,y) \in \mathbb{R}^2 : a < x < b,\ c < y < d\}$. Suppose that the partial derivatives of f exist and are bounded in S. Then show that f is continuous in S.*

Solution : Let $|f_x(x,y)| \leq M$ and $|f_y(x,y)| \leq M$ for all $(x,y) \in S$. Then

$$\begin{aligned} f(x+h,y+k) - f(x,y) &= f(x+h,y+k) - f(x+h,y) + f(x+h,y) - f(x,y) \\ &= kf_y(x+h, y+\theta_1 k) + hf_x(x+\theta_2 h, y), \text{ (for some } \theta_1, \theta_2 \in \mathbb{R}, \text{ by the MVT).} \end{aligned}$$

Hence, $| f(x+h,y+k) - f(x,y) | \leq M(|h| + |k|) \leq 2M\sqrt{h^2+k^2}$.

Hence, for $\epsilon > 0$, choose $\delta = \frac{\epsilon}{2M}$ or use the sequential argument to show that the function is continuous. □

It is clear from the previous example that the concept of differentiability of a function of several variables should be stronger than mere existence of partial derivatives of the function.

Differentiability : When $f : \mathbb{R} \to \mathbb{R}, x \in \mathbb{R}$ we define

$$f^{'}(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \qquad (*)$$

provided the limit exists. In case $f : \mathbb{R}^3 \to \mathbb{R}$ and $X = (x_1, x_2, x_3) \in \mathbb{R}^3$ the above definition of the differentiability of functions of one variable (*) cannot be generalized as we cannot divide by an element of $\mathbb{R}^3$. So, in order to define the concept of differentiability, what we do is that we rearrange the above definition (*) to a form which can be generalized.

Let $f : \mathbb{R} \to \mathbb{R}$. Then f is differentiable at x if and only if there exists $\alpha \in \mathbb{R}$ such that

$$\frac{| f(x+h) - f(x) - \alpha \cdot h |}{| h |} \to 0 \quad as \quad h \to 0.$$

When f is differentiable at x, α has to be $f^{'}(x)$. We generalize this definition to the functions of several variables.

Definition : Let $f : \mathbb{R}^3 \to \mathbb{R}$ and $X = (x_1, x_2, x_3)$. We say that f is *differentiable* at X if there exists $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3$ such that the error function

$$\varepsilon(H) = \frac{f(X + H) - f(X) - \alpha \cdot H}{\| H \|}$$

tends to 0 as $H \to 0$.

In the above definition $\alpha \cdot H$ is the scalar product. Note that the derivative $f'(X) = (\alpha_1, \alpha_2, \alpha_3)$.

Theorem 26.1: Let $f : \mathbb{R}^3 \to \mathbb{R}$, $X \in \mathbb{R}^3$. If f is differentiable at X then f is continuous at X.

Proof : Suppose f is differentiable at X. Then there exists $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3$ such that

$$\mid f(X + H) - f(X) - \alpha \cdot H \mid = \| H \| \, \varepsilon(H) \text{ and } \varepsilon(H) \to 0 \text{ as } H \to 0.$$

Hence

$$\mid f(X + H) - f(X) \mid \leq \| H \| \, (\sum_{i=1}^{3} \mid \alpha_i \mid) + \| H \| \, \varepsilon(H)$$

and $\varepsilon(H) \to 0$ as $H \to 0$. Therefore $f(X+H) \to f(X)$ as $H \to 0$. This proves that f is continuous at X. □

How do we verify that a given function is differentiable at a point in $\mathbb{R}^3$? The following result helps us to answer this question.

Theorem 26.2: Suppose f is differentiable at X. Then the partial derivatives $\frac{\partial f}{\partial x} \mid_X, \frac{\partial f}{\partial y} \mid_X$ and $\frac{\partial f}{\partial z} \mid_X$ exist and the derivative

$$f^{'}(X) = (\alpha_1, \alpha_2, \alpha_3) = (\frac{\partial f}{\partial x} \mid_X, \frac{\partial f}{\partial y} \mid_X, \frac{\partial f}{\partial z} \mid_X).$$

Proof : Suppose f is differentiable at X and $f'(X) = (\alpha_1, \alpha_2, \alpha_3)$. Then by taking $H = (t, 0, 0)$, we have

$$\varepsilon(H) = \frac{f(X + H) - f(X) - \alpha_1 t}{|t|} \to 0 \;\; as \;\; t \to 0, \;\; i.e., \;\; \frac{f(X + H) - f(X) - \alpha_1 t}{t} \to 0$$

This implies that $\alpha_1 = \frac{\partial f}{\partial x} \mid_X$. Similarly we can show that $\alpha_2 = \frac{\partial f}{\partial y} \mid_X$ and $\alpha_3 = \frac{\partial f}{\partial x} \mid_X$. □

Example 3 : Let

$$\begin{aligned} f(x, y) &= xy\frac{x^2 - y^2}{x^2 + y^2} \;\; at \;\; (x, y) \neq (0, 0) \\ &= 0 \;\; at \;\; (0, 0) \end{aligned}$$

To verify that f is differentiable at $(0, 0)$, let us choose $\alpha = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}) \mid_{(0,0)}$ and verify that $\epsilon(H) \to 0$ as $H = (h, k) \to 0$. In this case $\alpha = (0, 0)$ and

$$\mid \varepsilon(H) \mid = \mid \frac{f(0 + H) - f(0) - (0, 0) \cdot H}{\| H \|} \mid \leq \mid \frac{hk}{\sqrt{h^2 + k^2}} \mid \leq \sqrt{h^2 + k^2} \to 0 \;\; as \;\; H \to 0.$$

Hence f is differentiable at $(0, 0)$. □

Example 2 illustrates that the partial derivatives of a function at a point may exist but the function need not be differentiable at that point. The previous theorem says that if the function is differentiable at X then the derivative $f'(X)$ can be expressed in terms of the partial derivatives of f at X. Since finding partial derivatives is easy because they are based on one variable and it is related to the derivative, one naturally asks the following question: Under what additional assumptions on the partial derivatives the function becomes differentiable. The following criterion answer this question.

Theorem 26.3: If $f : \mathbb{R}^3 \to \mathbb{R}$ is such that all its partial derivatives exist in a neighborhood of X_0 and continuous at X_0 then f is differentiable at X_0.

We omit the proof of this result. We will see in a tutorial class that the converse of the previous result is not true.

Chain Rule: We have seen that the chain rule which deals with derivative of a function of a function is very useful in one variable calculus. In order to derive a similar rule for functions of several variables we need the following theorem called **Increment Theorem**. For simplicity we will state this theorem only for two variables.

We will employ the notation $f_x = \frac{\partial f}{\partial x}$ and $f_y = \frac{\partial f}{\partial x}$.

Theorem 26.4: Let $f(x, y)$ be differentiable at (x_0, y_0). Then we have

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) \;=\; f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y$$

where $\varepsilon_1(\Delta x, \Delta y), \varepsilon_2(\Delta x, \Delta y) \to 0$ as $\Delta x \to 0$ and $\Delta y \to 0$.

Proof (*): Let $H = (\Delta x, \Delta y)$. Since the function is differentiable at (x_0, y_0), we have

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \parallel H \parallel \varepsilon(H), \varepsilon(H) \to 0 \text{ as } H \to 0.$$

We have to show that $\parallel H \parallel \varepsilon(H) \;=\; \varepsilon_1 \Delta x + \varepsilon_2 \Delta y$ for some functions ε_1 and ε_2. Note that

$$\varepsilon(H) \parallel H \parallel = \frac{\varepsilon(H)}{\parallel H \parallel}(\Delta x^2 + \Delta y^2) \;=\; (\Delta x \frac{\varepsilon(H)}{\parallel H \parallel})\Delta x + (\Delta y \frac{\varepsilon(H)}{\parallel H \parallel})\Delta y.$$

Define $\varepsilon_1(H) = \Delta x \frac{\varepsilon(H)}{\|H\|}$ and $\varepsilon_2(H) = \Delta y \frac{\varepsilon(H)}{\|H\|}$. Note that

$$\mid \varepsilon_1(H) \mid = \mid \Delta x \frac{\varepsilon(H)}{\parallel H \parallel} \mid \le \mid \varepsilon(H) \mid \to 0 \text{ as } \; H \to 0.$$

Similarly we can show that $\varepsilon_2(H) \to 0$ as $H \to 0$. This proves the result. □

In the next result we present the chain rule.

Theorem 26.5: Let $f(x, y)$ be differentiable (or f has continuous partial derivatives) and if $x = x(t), y = y(t)$ are differentiable functions on t, then the function $w = f(x(t), y(t))$ is differentiable at t and

$$\frac{df}{dt} = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t), \;\; \text{i.e.,} \;\; \frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

Proof : By increment theorem we have

$$\Delta f \;=\; f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y, \quad \varepsilon_1, \varepsilon_2 \to 0 \;\text{ as }\; \Delta x, \Delta y \to 0$$

This implies that

$$\frac{\Delta f}{\Delta t} = f_x \frac{\Delta x}{\Delta t} + f_y \frac{\Delta y}{\Delta t} + \varepsilon_1 \frac{\Delta x}{\Delta t} + \varepsilon_2 \frac{\Delta y}{\Delta t}.$$

Allow $\Delta t \to 0$, which implies that $\varepsilon_1, \varepsilon_2 \to 0$ because $\Delta x, \Delta y \to 0$. Therefore, we get $\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$. □