

Lecture 10

Example 10.0.1 (Exercise)

Take the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ defined as,

$$f(x, y) = \begin{cases} \frac{x^3 y}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

- (i) Prove that, f is continuous at $(0, 0)$.
- (ii) Find $(D_u f)(0, 0) \forall u$.
- (iii) Prove that, f is not differentiable.

10.1 Extension of MVT to Several Variables

In Analysis I, you learned about the Mean Value Theorem (MVT) for functions of a single variable. Now, we extend this concept to several variables in the context of multivariable calculus.

Theorem 10.1.1 (Multivariate MVT)

Let $\mathcal{O}_n \subseteq \mathbb{R}^n$ be an open and convex set, and let $f : \mathcal{O}_n \rightarrow \mathbb{R}$ be a differentiable function. For any two points $a, b \in \mathcal{O}_n$ define the line segment

$$L_{a,b} := \{tb + (1-t)a : t \in [0, 1]\}$$

Then, there exists a point $c \in L_{a,b}$ such that,

$$f(b) - f(a) = (\nabla f)(c) \cdot (b - a) = \langle f_{x_1}(c), \dots, f_{x_n}(c) \rangle \cdot \langle (b_1 - a_1), \dots, (b_n - a_n) \rangle$$

Proof. We consider the function $\eta : [0, 1] \rightarrow \mathcal{O}_n$

$$\begin{array}{ccccc} [0, 1] & \xrightarrow{\eta} & \mathcal{O}_n & \xrightarrow{f} & \mathbb{R} \\ & \searrow & & \nearrow & \\ & & f \circ \eta & & \end{array}$$

defined by $\eta(t) = (1-t)a + tb$. This function is differentiable, and its derivative is

$$\eta'(t) = \begin{pmatrix} b_1 - a_1 \\ \vdots \\ b_n - a_n \end{pmatrix}$$

By applying the standard Mean Value Theorem to the composition $f \circ \eta$, there exists $t_0 \in (0, 1)$ such that,

$$f(\eta(1)) - f(\eta(0)) = (f \circ \eta)'(t_0) = Df(\eta(t_0)) \cdot D\eta(t_0)$$

Expanding the dot product, we have,

$$f(b) - f(a) = \begin{pmatrix} f_{x_1}(\eta(t_0)) & f_{x_2}(\eta(t_0)) & \cdots & f_{x_n}(\eta(t_0)) \end{pmatrix} \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \\ \vdots \\ b_n - a_n \end{pmatrix}$$

Simplifying further, we obtain,

$$f(b) - f(a) = \langle f_{x_1}(\eta(t_0)), \dots, f_{x_n}(\eta(t_0)) \rangle \cdot \langle (b_1 - a_1), \dots, (b_n - a_n) \rangle$$

This expression can be rewritten as,

$$f(b) - f(a) = (\nabla f)(\eta(t_0)) \cdot (b - a)$$

Hence, there exists $c = \eta(t_0) \in L_{a,b}$ such that $f(b) - f(a) = (\nabla f)(c) \cdot (b - a)$. $\square$

10.2 More Partial and Chain Rules

In this section, we further explore the chain rule for differentiable functions of several variables. Consider two functions f and g as following,

$$\mathcal{O}_n \xrightarrow{f} \mathcal{O}_m \xrightarrow{g} \mathbb{R}^p$$

Assuming that f is differentiable at $a \in \mathcal{O}_n$ and g is differentiable at $b = f(a) \in \mathcal{O}_m$ the chain rule states that the derivative of the composite function $g \circ f$ is given by,

$$\underbrace{D(g \circ f)(a)}_{\mathbb{R}^n \rightarrow \mathbb{R}^p} = (Dg)(f(a)) \cdot (Df)(a) \\ = \underbrace{Dg(b)}_{\mathbb{R}^m \rightarrow \mathbb{R}^p} \cdot \underbrace{Df(a)}_{\mathbb{R}^n \rightarrow \mathbb{R}^m}$$

This can be expressed in matrix form as,

$$J_{g \circ f}(a)_{p \times n} = J_g(f(a))_{p \times m} \cdot J_f(a)_{m \times n} \quad (10.1)$$

Moreover, if we consider the function components in each individual coordinates as

- $g \circ f = ((g \circ f)_1, (g \circ f)_2, \dots, (g \circ f)_p)$
- $g = (g_1, g_2, \dots, g_p)$
- $f = (f_1, f_2, \dots, f_m)$.

Then, the $(i, j)^{\text{th}}$ entry of both sides of (10.1) would become,

$$\frac{\partial (g \circ f)_i}{\partial x_j}(a) = \begin{pmatrix} \frac{\partial g_i}{\partial y_1}(f(a)) & \frac{\partial g_i}{\partial y_2}(f(a)) & \cdots & \frac{\partial g_i}{\partial y_m}(f(a)) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_j}(a) \\ \frac{\partial f_2}{\partial x_j}(a) \\ \vdots \\ \frac{\partial f_m}{\partial x_j}(a) \end{pmatrix}$$

which is our familiar chain rule for partial derivatives,

$$\frac{\partial (g \circ f)_i}{\partial x_j}(a) = \sum_{k=1}^m \frac{\partial g_i}{\partial y_k}(b) \cdot \frac{\partial f_k}{\partial x_j}(a)$$

where, $b = f(a)$.

Furthermore, if we define $y_k = f_k(x_1, x_2, \dots, x_n)$ and $z_i = g_i(y_1, y_2, \dots, y_m)$ we can express the chain rule for partial derivatives as,

$$\frac{\partial z_i}{\partial x_j}(x) = \sum_{k=1}^m \frac{\partial z_i}{\partial y_k}(y) \cdot \frac{\partial y_k}{\partial x_j}(x)$$

Remark. In addition, the chain rule can be applied in the context of a composition map with respect to a parameter t . For functions $f : \mathcal{O}_n \rightarrow \mathbb{R}$, $\eta : \mathcal{O}_1 \rightarrow \mathcal{O}_n$ and $z = f \circ \eta$ as shown below,

$$\begin{array}{ccccc} \mathcal{O}_1 & \xrightarrow{\eta} & \mathcal{O}_n & \xrightarrow{f} & \mathbb{R} \\ \\ t & \xrightarrow{\eta} & (x_1(t), \dots, x_n(t)) & \xrightarrow{f} & f(x_1(t), \dots, x_n(t)) \\ & & \searrow & \nearrow & \\ & & & z = f \circ \eta & \end{array}$$

the chain rule states,

$$\frac{dz}{dt} = \sum_{k=1}^n \frac{\partial f}{\partial x_k} \cdot \frac{dx_k}{dt}$$

If we treat f as a function of t , the same can be written as,

$$\frac{df}{dt} = \sum_{k=1}^n \frac{\partial f}{\partial x_k} \cdot \frac{dx_k}{dt}$$