Lecture 9: The Chain Rule & The Mean value Theorem

October 26, 2018

Sunil Kumar Gauttam

Department of Mathematics, LNMIIT

The Chain Rule for functions of a single variable says that when $f(x)$ is a differentiable function of x and $x = g(t)$ is a differentiable function of t, $w = f \circ g$ is a differentiable function of t and

$$\frac{dw}{dt}(t) = \frac{df}{dx}(g(t))\frac{dg}{dt}(t)$$

For functions of two variables the Chain Rule has three forms, simply for the reason that there are three types of composites possible.

1. If $f(x, y)$ is differentiable and $g(t)$ is differentiable, then the composite $w(x, y) = g(f(x, y))$ is a differentiable function of (x, y), and

$$\frac{\partial w}{\partial x}(x, y) = \frac{dg}{dt}(f(x, y))\frac{\partial f}{\partial x}(x, y) \quad \text{and} \quad \frac{\partial w}{\partial y}(x, y) = \frac{dg}{dt}(f(x, y))\frac{\partial f}{\partial y}(x, y)$$

Example 9.1 *Consider $f : \mathbb{R}^2 \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ defined by $f(x, y) := xy$ for $(x, y) \in \mathbb{R}^2$ and $g(t) = \sin t$ for $t \in \mathbb{R}$. By chain rule the composite function $w : \mathbb{R}^2 \to \mathbb{R}$ given by $w(x, y) = (g \circ f)(x, y) = \sin(xy)$ is differentiable at every point of $\mathbb{R}^2$ and*

$$\frac{\partial w}{\partial x}(x, y) = \cos(xy)\ y \quad \frac{\partial w}{\partial y}(x, y) = \cos(xy)\ x$$

An alternative approach to the Example 9.1, is to substitute the formulas for f in terms of x and y directly into the formula for g. By doing so, we express

$$w(x, y) = \sin(xy)$$

so that we may calculate the partial derivative directly. In spite of the fact that the Chain Rule is not absolutely necessary for calculations such as in Example 9.1 , we will see that the Chain Rule is essential for certain theoretical developments such as mean value theorem.

2. If $f(x, y)$ is differentiable and if $x = x(t), y = y(t)$ are is differentiable, then $z = f((x(t), y(t))$ is a is differentiable function of t, and

$$\frac{dz}{dt}(t) = \frac{\partial f}{\partial x}(x(t), y(t))\frac{dx}{dt}(t) + \frac{\partial f}{\partial y}(x(t), y(t))\frac{dy}{dt}(t)$$

3. If $f(x,y)$ is differentiable and if $x = x(u,v), y = y(u,v)$ are differentiable, then $z(u,v) = f(x(u,v), y(u,v))$ is a differentiable function of (u,v), and

$$\frac{\partial z}{\partial u}(u,v) = \frac{\partial f}{\partial x}(x(u,v), y(u,v))\frac{\partial x}{\partial u}(u,v) + \frac{\partial f}{\partial y}(x(u,v), y(u,v))\frac{\partial y}{\partial u}(u,v)$$
$$\text{and} \quad \frac{\partial z}{\partial v}(u,v) = \frac{\partial f}{\partial x}(x(u,v), y(u,v))\frac{\partial x}{\partial v}(u,v) + \frac{\partial f}{\partial y}(x(u,v), y(u,v))\frac{\partial y}{\partial v}(u,v)$$

Theorem 9.2 (MVT for functions of two variables) *Let $D \subseteq \mathbb{R}^2$ be a closed disk and suppose $f : D \to \mathbb{R}$ is differentiable. Given any distinct points $A = (x_0, y_0)$ and $B = (x_1, y_1)$ in D, there is $C = (c, d)$ lying on the line segment joining A and B with $C \neq A, C \neq B$ such that*

$$f(B) - f(A) = f'(C) \cdot (B - A).$$

In other words,

$$f(x_1, y_1) - f(x_0, y_0) = (x_1 - x_0, y_1 - y_0) \cdot \nabla f(c, d) = (x_1 - x_0)f_x(c, d) + (y_1 - y_0)f_y(c, d).$$

Proof: Define $\phi : [0, 1] \to \mathbb{R}$ by

$$\phi(t) := f(x_0 + (x_1 - x_0)t, y_0 + (y_1 - y_0)t).$$

Since f is differentiable on D and $x(t) = x_0+(x_1-x_0)t, y(t) = y_0+(y_1-y_0)t$ are differentiable function on $\mathbb{R}$. Hence ϕ is differentiable on $[0, 1]$ (hence continuous) and by chain rule

$$\phi'(t) = \nabla f(x(t), y(t)) \cdot (x'(t), y'(t))$$

Applying MVT for ϕ, there exist $s \in (0, 1)$ such that

$$\phi(1) - \phi(0) = \phi'(s)(1 - 0).$$

That is

$$f(x_1, y_1) - f(x_0, y_0) = (x_1 - x_0)f_x(c, d) + (y_1 - y_0)f_y(c, d),$$

where $c = x_0 + (x_1 - x_0)s, d = y_0 + (y_1 - y_0)s$. ■

Corollary 9.3 *Let $f : \mathbb{R}^2 \to \mathbb{R}$ such that both f_x and f_y vanish identically on the plane, then f is constant on $\mathbb{R}^2$.*

Proof: Since f_x and f_y exists throughout the plane and are zero, hence f_x and f_y are continuous everywhere. Therefore f is differentiable on $\mathbb{R}^2$. Let (x_0, y_0) be any point of $\mathbb{R}^2$. Since $f_x = f_y = 0$ on $\mathbb{R}^2$, by Theorem 9.2, for any $(x_1, y_1) \in D$ with $(x_1, y_1) \neq (x_0, y_0)$ we have $f(x_1, y_1) - f(x_0, y_0) = 0$, that is, $f(x_1, y_1) = f(x_0, y_0)$. Thus, f is a constant function. ■