**MULTIVARIABLE CALCULUS**
**LECTURE 23**

## 1. Chain Rule

The general Chain Rule with two variables We the following general Chain Rule is needed to find derivatives of composite functions in the form $z = f(x(t), y(t))$ or $z = f(x(s,t), y(s,t))$. We begin with functions of the first type.

**Theorem 1.1. (The Chain Rule)** *The $t$-derivative of the composite function $z = f(x(t), y(t))$ is*

$$\frac{d}{dt}[f(x(t), y(t))] = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t) \tag{1}$$

*assuming $f_x$ and $f_y$ are continuous.*

We assume in this theorem and its applications that $x = x(t)$ and $y = y(t)$ have first derivatives at $t$ and that $z = f(x, y)$ has continuous first-order derivatives in an open disc centered at $(x(t), y(t))$. Equation (1) can be read as the following statement: the $t$-derivative of the composite function equals the $x$-derivative of the outer function $z = f(x, y)$ at the point $(x(t), y(t))$ multiplied by the $t$-derivative of the inner function $x = x(t)$, plus the y-derivative of the outer function at $(x(t), y(t))$ multiplied by the $t$-derivative of the inner function $y = y(t)$.

We fix $t$ and set $(x, y) = (x(t), y(t))$. We consider nonzero $\Delta t$ so small that $(x(t + \Delta t), y(t + \Delta t))$ is in the circle where $f$ has continuous first derivatives and set $\Delta x = x(t + \Delta t) - x(t)$ and $\Delta y = y(t + \Delta t) - y(t)$. Then, by the definition of the derivative,

$$\begin{aligned}
\frac{d}{dt}[f(x(t), y(t))] &= \lim_{\Delta t \to 0} \frac{f(x(t + \Delta t), y(t + \Delta t)) - f(x(t), y(t))}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{f(x(t) + \Delta x, y(t) + \Delta y) - f(x(t), y(t))}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{f(x + \Delta x, y + \Delta y) - f(x(t), y(t))}{\Delta t} \qquad (2)
\end{aligned}$$

We express the change $f(x(t) + \Delta x, y(t) + \Delta y) - f(x(t), y(t))$ in the value of $z = f(x, y)$ from $(x, y)$ to $x + \Delta x, y + \Delta y$ as the change in the $x$-direction from $(x, y)$ to $(x + \Delta x, y)$ plus the change in the $y$-direction from $(x + \Delta x, y)$ to $(x + \Delta x, y + \Delta y)$.

(3)
$$f(x + \Delta x, y + \Delta y) - f(x, y) = [f(x + \Delta x, y) - f(x, y)] + [f(x + \Delta x, y + \Delta y) - f(x + \Delta x, y)].$$

(Notice that the terms $f(x + \Delta x, y)$ and $-f(x + \Delta x, y)$ on the right side of (3) cancel to give the left side).
We can apply the Lagrange Mean Value Theorem to the expression in the first set of square brackets on the right of (3) by keeping $y$ constant and to the expression in the second set

of square brackets where keeping $x$ constant. By using LMVT, we conclude that there is a number $c_1 \in (x, x+\Delta x)$ and a number $c_2 \in (y, y+\Delta y)$ such that

$$\begin{aligned} &f(x+\Delta x, y) - f(x,y) = f_x(c_1, y)\Delta x \\ &f(x+\Delta x, y+\Delta y) - f(x+\Delta x, y) = f_y(x+\Delta x, c_2)\Delta y. \end{aligned} \tag{4}$$

We combine equations (3) and (4) and divide by $\Delta t$ to obtain

$$\frac{f(x+\Delta x, y+\Delta y) - f(x(t), y(t))}{\Delta t} = f_x(c_1, y)\frac{\Delta x}{\Delta t} + f_y(x+\Delta x, c_2)\frac{\Delta y}{\Delta t}. \tag{5}$$

The functions $x = x(t)$ and $y = y(t)$ are continuous at $t$ because they have derivatives at that point. Consequently, as $\Delta t \to 0$, the numbers $\Delta x = x(t+\Delta t) - x(t)$ and $\Delta y = y(t+\Delta t) - y(t)$ both tend to zero. Because the partial derivatives of $f$ are continuous, so $f_x(c_1, y) \to f_x(x, y)$ and $f_y(x+\Delta x, c_2) \to f_y(x, y)$ as $\Delta t \to 0$.

Moreover $\frac{\Delta x}{\Delta t} \to\to x'(t)$ as $\Delta t \to 0$. And $\frac{\Delta y}{\Delta t} \to\to y'(t)$ as $\Delta t \to 0$.

So equation (2) and (5) gives

$$\frac{d}{dt}[f(x(t), y(t))] = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t)$$

The above theorem, can be applied to find the s- and t-derivatives of a function of the form $z = f(x(s,t), y(s,t))$ because in taking the derivative with respect to $s$ or $t$, the other variable is constant. We obtain the following :

**Theorem 1.2. (The Chain Rule)** *The s- and t-derivatives of the composite function $z = f(x(s,t), y(s,t))$ are*

$$\begin{aligned} &\frac{d}{ds}[f(x(s,t), y(s,t))] = f_x(x(s,t), y(s,t))x_s(s,t) + f_y(x(s,t), y(s,t))y_s(s,t), \\ &\frac{d}{dt}[f(x(s,t), y(s,t))] = f_x(x(s,t), y(s,t))x_t(s,t) + f_y(x(s,t), y(s,t))y_t(s,t), \end{aligned} \tag{6}$$

We assume in this theorem and its applications that the functions involved have continuous first derivatives in the open sets where they are considered. Formulas (6) are easier to remember without the values of the variables in the form,

$$\begin{aligned} f_s &= f_x x_s + f_y y_s \\ f_t &= f_x x_t + f_y y_t. \end{aligned}$$

**Example :** Let $z = \log(u^2+v^2)$, $u(x,y) = e^{x+y^2}$, $v(x,y) = x^2+y$. Then

- $\dfrac{\partial z}{\partial u} = \dfrac{2u}{u^2+v^2}$, $\dfrac{\partial z}{\partial v} = \dfrac{2v}{u^2+v^2}$, $\dfrac{\partial u}{\partial x} = e^{x^2+y}$, $\dfrac{\partial v}{\partial x} = 2x$.
- $\dfrac{\partial z}{\partial x} = \dfrac{\partial z}{\partial u}\dfrac{\partial u}{\partial x} + \dfrac{\partial z}{\partial v}\dfrac{\partial v}{\partial x} = \dfrac{2u}{u^2+v^2}e^{x+y^2} + \dfrac{2xv}{u^2+v^2}$. Similarly we can compute $\dfrac{\partial z}{\partial y}$.

1.1. **Derivative of implicitly defined function.** Let $y = y(x)$ be defined as $F(x, y(x)) = 0$, where $F, F_x, F_y$ are continuous at $(x_0, y_0)$ and $F_y(x_0, y_0) \neq 0$. Then

$$\frac{dy}{dx} = -\frac{F_x}{F_y} \quad \text{at } (x_0, y_0).$$

*Proof.* Let us define $g(x) = F(x, y(x))$. Then differentiate with respect to $x$, we obtain using Chain rule

$$\frac{dg(x)}{dx} = g'(x) = F_x.\frac{dx}{dx} + F_y\frac{dy(x)}{dx} = 0.$$

This implies

$$\frac{dy(x)}{dx} = -\frac{F_x(x_0, y_0)}{F_y(x_0, y_0)} \quad \text{at } (x_0, y_0).$$

□

**Example :**The function $y(x)$ defined implicitly as $e^y - e^x + xy = 0$. Let $F(x, y) = e^y - e^x + xy$. Then

$$\frac{\partial F}{\partial x} = -e^x + y, \quad \frac{\partial F}{\partial y} := e^y + x.$$

Then

$$\frac{dy(x)}{dx} = \frac{e^x - y}{e^y + x}.$$

1.2. **Gradient.** Let us recall the definition of differentiable function. A function $f$ is said to be differentiable at the point $(a, b)$ if there exist and $\varepsilon_1$ and $\varepsilon_2$ such that

$$f(a + h, b + k) = f(a, b) = (\frac{\partial f}{\partial x}(a, b), \frac{\partial f}{\partial y}(a, b)).(h, k) + h\varepsilon_1 + k\varepsilon_2,$$

where $\varepsilon_1 \to 0$ and $\varepsilon_2 \to 0$, as $(h, k) \to (0, 0)$.

**Definition 1.1.** Gradient of $f$ at a point $(a, b)$ is defined to be a vector in $\mathbb{R}^2$,( written)

$$\nabla f(a, b) := \left( (\frac{\partial f}{\partial x}(a, b), \frac{\partial f}{\partial y}(a, b)) \right) \in \mathbb{R}^2.$$

**Remark 1.1.** The gradient vector has lot of geometric significance. Moreover it is evident from the definition that the gradient vector may exists even even when the function is not differentiable at some point.

If the function is differentiable at some point then we have the following proposition.

**Proposition 1.3.** *If $f(x, y)$ is differentiable, then the directional derivative in the direction $\vec{p}$ at $(a, b)$ is*

$$D_{\vec{p}} f(a, b) = \nabla f(a, b).\vec{p}.$$

*Proof.* Let $\vec{p} = p_1\hat{i} + p_2\hat{j}$. Using the definition of directional derivative we have

$$\begin{aligned} D_{\vec{p}} f(a, b) &= \lim_{s \to 0} \frac{f(a + sp_1, b + sp_2) - f(a, b)}{s} \\ &= \lim_{s \to 0} \frac{f(x(s), y(s)) - f(x(0), y(0))}{s} \\ &= \frac{d}{ds} f(x(s), y(s)) \quad \text{at } s = 0, \end{aligned}$$

where $x(s) = a + sp_1$ and $y(s) = b + sp_2$ Now applying chain rule we have at $s = 0$

$$\begin{aligned}\frac{d}{ds}\, f(x(s), y(s)) &= \frac{\partial f}{\partial x}(a,b)\frac{dx}{ds} + \frac{\partial f}{\partial y}(a,b)\frac{dy}{ds}\\ &= \left( (\frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b)) \right).(p_1, p_2)\\ &= \nabla f(a,b).(p_1,p_2).\end{aligned}$$

Hence the result follows. □

**Remark 1.2.** Geometrical interpretation of gradient. Suppose $f$ is differentiable at $(a,b)$. Then from the previous proposition we obtain

$$D_{\vec{p}} f(a,b) = \nabla f(a,b).\vec{p} = \underbrace{|\nabla f(a,b)|\cos\theta}_{\text{max when} \cos\theta = 1},$$

where $\vec{p}$ is an unit vector. Therefore the maximum when $\vec{p}$ is in the direction of $\nabla f(a,b)$. So $\vec{p} = \dfrac{\nabla f(a,b)}{|\nabla f(a,b)|}$ and

$$D_{\vec{p}} f(a,b) = |\nabla f(a,b)|.$$

In particular, $f$ increases most rapidly in the direction of $\nabla f(a,b)$. Moreover $f$ decreases rapidly in the direction of $-\nabla f$.

**Remark 1.3.** The formula $D_{\vec{p}} f(a,b) = \nabla f(a,b).\vec{p}$ can still hold even when the function is not differentiable.

Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x,y) := \begin{cases} \dfrac{x^2 y\sqrt{|y|}}{x^4 + y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = 0 \end{cases}$$

One can easily see that any directional derivative of $f$ at $(0,0)$ is 0 and also $f_x(0,0) = f_y(0,0) = 0$. So the equality holds trivially. Although the function is not differentiable. Indeed,

$$\frac{\Delta f - df}{\rho} = \frac{h^2 k\sqrt{|k|}|}{h^4 + k^2}.\frac{1}{\sqrt{h^2 + k^2}}$$

substitute $h = r\cos\theta$ and $h = r\sin\theta$, we get

$$\frac{h^2 k\sqrt{|k|}|}{h^4 + k^2}.\frac{1}{\sqrt{h^2 + k^2}} = \sqrt{r}\frac{\cos^2\theta\sin\theta\sqrt{|\sin\theta|}}{r^2\cos^4\theta + \sin^2\theta}.$$

Further we choose a path $r = \dfrac{\sin\theta}{\cos\theta}$, from the above we obtain

$$\frac{h^2 k\sqrt{|k|}|}{h^4 + k^2}.\frac{1}{\sqrt{h^2 + k^2}} = |\sin\theta|\frac{\cos\theta}{\sin^2\theta(1 + \cos^2\theta)} \to \infty, \quad \text{as } \theta \to 0.$$

Hence NOT differentiable at $(0,0)$.