SE102:Multivariable Calculus

Hyosang Kang[1]

[1]Division of Mathematics
School of Interdisciplinary Studies
DGIST

Lecture 03
The Chain Rule

Theorem (The chain rule)

Let $f(x, y)$ be a two variable function and $c(t) = (x(t), y(t))$ be a parametrization of a curve in $\mathbf{R}^2$ which is differentiable at t_0. If $f(x, y)$ is differentiable at $(x_0, y_0) = (x(t_0), y(t_0))$, then the composition

$$F(t) = (f \circ c)(t) = f(x(t), y(t))$$

is also differentiable at t_0 and its differential is

$$F'(t_0) = f_x(x_0, y_0)x'(t_0) + f_y(x_0, y_0)y'(t_0)$$

Proof.

Let us define $g_1(t), g_2(t)$ as follows.

$$g_1(t) = \frac{x(t) - x_0 - x'(t_0)(t - t_0)}{t - t_0}$$

$$g_2(t) = \frac{y(t) - y_0 - y'(t_0)(t - t_0)}{t - t_0}$$

Since $x(t)$, $y(t)$ are differentiable at t_0,

$$\lim_{t \to t_0} g_1(t) = \lim_{t \to t_0} g_2(t) = 0.$$

Thus $g_1(t)$, $g_2(t)$ are continuous. Define $F(x, y)$ as

$$F(x, y) = \frac{f(x, y) - L(x, y)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}}.$$

Proof.

Since we assumed that $f(x, y)$ is differentiable at (x_0, y_0),

$$\lim_{(x,y)\to(x_0,y_0)} F(x, y) = 0.$$

Thus $F(x, y)$ is continuous on $\mathbf{R}^2$. Note that

$$x(t) - x_0 = (x'(t_0) + g_1(t))(t - t_0),$$

$$y(t) - y_0 = (y'(t_0) + g_2(t))(t - t_0),$$

and we can rewrite the definition of $F(x, y)$ as

$$\begin{aligned} f(x, y) - f(x_0, y_0) = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\ + F(x, y)\sqrt{(x - x_0)^2 + (y - y_0)^2}. \end{aligned}$$

Proof.

With further computation, we get

$$
\begin{aligned}
&\frac{f(x(t),y(t)) - f(x_0,y_0)}{t-t_0} \\
&\quad = \frac{1}{t-t_0}\Big(f_x(x_0,y_0)(x(t)-x_0) + f_y(x_0,y_0)(y(t)-y_0) \\
&\qquad\qquad + F(x(t),y(t))\sqrt{(x(t)-x_0)^2+(y(t)-y_0)^2}\Big) \\
&\quad = f_x(x_0,y_0)(x'(t_0)+g_1(t)) + f_y(x_0,y_0)(y'(t_0)+g_2(t)) \\
&\qquad\qquad + F(x(t),y(t))\sqrt{\left(\frac{x(t)-x_0}{t-t_0}\right)^2+\left(\frac{y(t)-y_0}{t-t_0}\right)^2}
\end{aligned}
$$

Proof.

Let us take the limit on both sides:

$$
\begin{aligned}
&\lim_{t\to t_0} \frac{f(x(t),y(t)) - f(x_0,y_0)}{t-t_0} \\
&\quad = f_x(x_0,y_0)x'(t_0) + f_y(x_0,y_0)y'(t_0) \\
&\qquad + f_x(x_0,y_0)\lim_{t\to t_0} g_1(t) + f_y(x_0,y_0)\lim_{t\to t_0} g_2(t) \\
&\qquad + \lim_{t\to t_0} F(x(t),y(t))\sqrt{\left(\lim_{t\to t_0}\frac{x(t)-x_0}{t-t_0}\right)^2 + \left(\lim_{t\to t_0}\frac{y(t)-y_0}{t-t_0}\right)^2} \\
&\quad = f_x(x_0,y_0)x'(t_0) + f_y(x_0,y_0)y'(t_0).
\end{aligned}
$$

This proves the theorem.

Remark

We can write the chain rule using the gradient of $f(x, y)$.

$$(f \circ c)'(t_0) = \nabla f(c(t_0)) \cdot c'(t_0)$$

The formula says that *the rate of change of* $f(x, y)$ *at* (x_0, y_0) *to the direction of* $c'(t_0)$ *is given by the inner product.*
Thus we can say that *the gradient measures how much* $f(x, y)$ *changes to the given direction.*

Remark

We can define the gradient $\nabla f(x_0, y_0)$ as a linear map $\nabla f(x_0, y_0) : \mathbf{R}^2 \to \mathbf{R}$ such that

$$\nabla f(x_0, y_0)(\mathbf{v} + \mathbf{w}) = \nabla f(x_0, y_0) \cdot \mathbf{v} + \nabla f(x_0, y_0) \cdot \mathbf{w}$$

for all $\mathbf{v}$, $\mathbf{w}$ in $\mathbf{R}^2$. We will see that this is a particular case of /emphdifferential.

Definition

Let $F : \mathbf{R}^2 \to \mathbf{R}^2$ be a two-variable function whose coordinates are defined as follows:

$$F(u, v) = (x(u, v), y(u, v))$$

We say F is continuous (differentiable, repectively) at (u_0, v_0) if all coordinate functions $x, y : \mathbf{R}^2 \to \mathbf{R}$ are continuous (differentiable, respectively) at (u_0, v_0).

Theorem

Let $\mathbf{X} : \mathbf{R}^2 \to \mathbf{R}^2$ be a function defined as

$$\mathbf{X}(u, v) = (x(u, v), y(u, v)).$$

Let us denote $x_0 = x(u_0, v_0)$, $y_0 = y(u_0, v_0)$. If X are differentiable at (u_0, v_0), then for a function $f : \mathbf{R}^2 \to \mathbf{R}$ which is differentiable at (x_0, y_0), the composite function $F = f \circ X$ is also differentiable at (u_0, v_0). Moreoever,

$$F_u(u_0, v_0) = f_x(x_0, y_0)x_u(u_0, v_0) + f_y(x_0, y_0)y_u(u_0, v_0),$$
$$F_v(u_0, v_0) = f_x(x_0, y_0)x_v(u_0, v_0) + f_y(x_0, y_0)y_v(u_0, v_0)$$

Proof.

If we assume that $F = f \circ X$ is differentiable, then by the chain rule that we proved earlier, we get

$$\begin{aligned} F_u(u_0, v_0) &= \frac{d}{du}\Big|_{u=u_0} f(x(u, v_0), y(u, v_0)) \\ &= f_x(x_0, y_0) x_u(u_0, v_0) + f_y(x_0, y_0) y_u(u_0, v_0), \\ F_v(u_0, v_0) &= \frac{d}{dv}\Big|_{v=v_0} f(x(u_0, v), y(u_0, v)) \\ &= f_x(x_0, y_0) x_v(u_0, v_0) + f_y(x_0, y_0) y_v(u_0, v_0) \end{aligned}$$

Let us define

$$\begin{aligned} S(u, v) = {} & F(u, v) - F(u_0, v_0) \\ & - \begin{bmatrix} f_x(x_0, y_0) & f_y(x_0, y_0) \end{bmatrix} \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} \end{aligned}$$

Proof.

It suffices to show that $\displaystyle\lim_{(u,v)\to(u_0,v_0)} \frac{S(u,v)}{\sqrt{(u-u_0)^2+(v-v_0)^2}} = 0.$

Let us define

$$H(u,v) = \frac{X(u,v) - X(u_0,v_0) - \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} \begin{bmatrix} u-u_0 \\ v-v_0 \end{bmatrix}}{\sqrt{(u-u_0)^2+(v-v_0)^2}}.$$

From the differentiability of X, we have $\displaystyle\lim_{(u,v)\to(u_0,v_0)} H(u,v) = 0.$

Note that $X(u,v) = \begin{bmatrix} x \\ y \end{bmatrix}$, $X(u_0,v_0) = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$. Thus

$$\begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} \begin{bmatrix} u-u_0 \\ v-v_0 \end{bmatrix} = \begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix} - H(u,v)\sqrt{(u-u_0)^2+(v-v_0)^2}$$

Proof.

By substituting above into the formula of $S(u, v)$, we get

$$S(u,v) = f(x,y) - f(x_0,y_0) - \begin{bmatrix} f_x(x_0,y_0) & f_y(x_0,y_0) \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} + \begin{bmatrix} f_x(x_0,y_0) & f_y(x_0,y_0) \end{bmatrix} H(u,v) \sqrt{(u-u_0)^2 + (v-v_0)^2}$$

Then,

$$\frac{S(u,v)}{\sqrt{(u-u_0)^2 + (v-v_0)^2}} = \frac{f(x,y) - f(x_0,y_0) - \begin{bmatrix} f_x(x_0,y_0) & f_y(x_0,y_0) \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}}{\sqrt{(u-u_0)^2 + (v-v_0)^2}} + \begin{bmatrix} f_x(x_0,y_0) & f_y(x_0,y_0) \end{bmatrix} H(u,v)$$

Proof.

By the differentiability of X, the second summand vanishes as $(u, v) \to (u_0, v_0)$. We can rewrite the first summand as multiple of two factors below:

$$\frac{f(x,y) - f(x_0,y_0) - \begin{bmatrix} f_x(x_0,y_0) & f_y(x_0,y_0) \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}}{\sqrt{(x-x_0)^2 + (y-y_0)^2}} \tag{1}$$

$$\frac{\sqrt{(x-x_0)^2 + (y-y_0)^2}}{\sqrt{(u-u_0)^2 + (v-v_0)^2}} \tag{2}$$

(**??**) vanishes as $(x, y) \to (x_0, y_0)$ due to the differentiability of f. Thus it suffices to prove that (**??**) is bounded. Note that $|X(u,v) - X(u_0,v_0)| = \sqrt{(x-x_0)^2 + (y-y_0)^2}$. Thus (**??**) satisfies

$$\frac{\sqrt{(x-x_0)^2 + (y-y_0)^2}}{\sqrt{(u-u_0)^2 + (v-v_0)^2}} \leq |H(u,v)| + \left| \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} \right|$$

Proof.

Again, $H(u,v)$ is vanishes as $(u,v) \to (u_0, v_0)$. By the inequality

$$\left| \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} \right| \le \sqrt{2} \max\{|x_u|, |x_v|, |y_u|, |y_v|\} \left| \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} \right|,$$

we are done.

Remark

We can write the chain rule using matrices. For $F(t) = (f \circ c)(t)$,

$$F'(t) = \begin{bmatrix} f_x & f_y \end{bmatrix} \cdot \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix}$$

For $F(u, v) = (f \circ X)(u, v)$,

$$\begin{bmatrix} F_u & F_u \end{bmatrix} = \begin{bmatrix} f_x & f_y \end{bmatrix} \cdot \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix}.$$

Note that the gradient of F and $f \circ X$ can be viewed as 1×2 matrices. Thus the last equation can be written as

$$\nabla (f \circ X) = \nabla f \cdot DX$$

Example

Let $T(r,\theta) = (r\cos\theta, r\sin\theta)$. Suppose that $F = f \circ T$. Then

$$F_r = f_x x_r + f_y y_r = f_x \cos\theta + f_y \sin\theta$$

$$F_\theta = f_x x_\theta + f_y y_\theta = -r f_x \sin\theta + r f_y \cos\theta$$

In matrix form,

$$\begin{bmatrix} F_r & F_\theta \end{bmatrix} = \begin{bmatrix} f_x & f_y \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}.$$

Definition

Let $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$ be a differentiable function. Then the **differential** at $\mathbf{a} \in \mathbf{R}^n$ is the $m \times n$ matrix $\mathbf{df}(\mathbf{a})$ defined as below:

$$\mathbf{df}(\mathbf{a}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{bmatrix}$$

Here, each f_i is the coordinate function for f:

$$f(\mathbf{x}) = (f_1(\mathbf{x}), \cdots, f_n(\mathbf{x})).$$

Theorem (Chain rule)

Let $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$ *be a differentiable function at* $\mathbf{a} \in \mathbf{R}^n$ *and* $\mathbf{g} : \mathbf{R}^m \to \mathbf{R}^l$ *be a differentiable function at* $\mathbf{f}(\mathbf{a}) \in \mathbf{R}^m$*. Then* $\mathbf{g} \circ \mathbf{f}$ *is a differentiable function at* $\mathbf{a}$*, and*

$$\mathbf{d}(\mathbf{g} \circ \mathbf{f})(\mathbf{a}) = \mathbf{dg}(\mathbf{f}(\mathbf{a}))\mathbf{df}(\mathbf{a})$$

Definition

Let S be a subset of $\mathbf{R}^n$. The **tangent space** $T_{\mathbf{a}}S$ at $\mathbf{a} \in S$ is the vector space consists of all *tangent* vectors of S at $\mathbf{a}$.

Remark

The tangent space $T_{\mathbf{a}}\mathbf{R}^n$ is a n-dimensional vector space $\mathbf{R}^n$. Sicne $\mathbf{df}(\mathbf{a})$ is a matrix, one can view the differential $\mathbf{df}(\mathbf{a})$ as a linear map $\mathbf{df}(\mathbf{a}) : \mathbf{R}^n \to \mathbf{R}^m$ as follows: for each n-dimensional vector $\mathbf{v} = (v_1, \cdots, v_n)$ in $T_{\mathbf{a}}\mathbf{R}^n$, the vector $\mathbf{df}(\mathbf{a})\mathbf{v}$ is a m-dimensional vector defined by

$$d\mathbf{f}(\mathbf{a})\mathbf{v} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

Example

Let $f(x, y, z) = x + 2y + 3z$. Let S be the graph of $z = xy$. The differential $df(p)$ at $p = (1, 1, 1)$ on S is

$$df(1,1,1) = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}.$$

The tangent plane of S at p is

$$x + y - z = 1$$

and the tangent space T_pS is spanned by

$$\mathbf{v}_1 = (1, 0, 1), \quad \mathbf{v}_2 = (0, 1, 1).$$

The parametric curve $c(t) = (t, 1, t)$ lies on S, Since $c(1) = p$ and $c'(1) = \mathbf{v}_1$, by the chain rule,

$$(f \circ c)'(1) = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 4$$

In other words, $df(p) \cdot \mathbf{v}_1 = 4$. Similarly, we can compute $df(p) \cdot \mathbf{v}_2 = 5$. Since given a tangent vector $\mathbf{v} = a\mathbf{v}_1 + b\mathbf{v}_2$ for suitable a, b, the rate of change of f in direction of $\mathbf{v}$ is

$$df(p) \cdot \mathbf{v} = 4a + 5b.$$

Problem

A cylinderical ice is melting in a room. When the radius of ice is 6(cm) and the height is 10(cm), the radius is decreasing at 0.1(cm/min) and the height is decreasing at 0.2(cm/min). How fast (cmˆ3/min) is the ice melting?

Problem

Let $f(x,y) = \sqrt{|xy|}$. Find the tangent plane to the graph at $(1,1,1)$.

Problem

Let $f(x, y) = (x^2 - y^2, 2xy)$.

1. Find the differential $df(1, 1)$.
2. Let $D = [1, 1 + \varepsilon] \times [1, 1 + \varepsilon]$. Compute the limit

$$\lim_{\varepsilon \to 0} \frac{\mathrm{area} f(D)}{\mathrm{area}(D)}$$

3. Find any relation between results in 1 and 2.

Problem

Let $f(x, y)$ be a function with continuous partial derivatives. Let $x = e^r \cos\theta$ and $y = e^r \sin\theta$. Show that

$$\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 = e^{-2r}\left(\left(\frac{\partial f}{\partial r}\right)^2 + \left(\frac{\partial f}{\partial \theta}\right)^2\right)$$

Problem

The *Laplacian* Δf of $f(x, y)$ is defined as

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Show that

$$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2}$$