Multivariable Calculus

MATH2007

1.3 The Chain Rule (Part 1)

WITS UNIVERSITY

School of Mathematics

Theorem (1.3.1, $\mathbb{R} \to \mathbb{R}$ Chain Rule)**.** Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$, then

$$\frac{d(f \circ g)}{dt}(t) = \left.\frac{df}{dt}\right|_{g(t)} \frac{dg}{dt}(t)$$

[i.e. $(f \circ g)'(t) = f'(g(t))g'(t)$.]

Proof. See tutorial Q1. $\square$

Example. Illustrate the chain rule for $f(x) = x^3$ and $g(x) = 1 + 2x$.

$$\frac{d}{dx}(f\circ g)(x) = f'(g(x))g'(x) = 3x^2\Big|_{x \to 1+2x} \cdot 2$$

$$= 3(1+2x)^2 \cdot 2 = 24x^2 + 24x + 6.$$

$$(f\circ g)(x) = x^3\Big|_{x\to 1+2x} = (1+2x)^3 = 8x^3 + 12x^2 + 6x + 1$$

$$\frac{d}{dx}(f\circ g)(x) = 24x^2 + 24x + 6 \quad \checkmark$$

Theorem (1.3.2, $\mathbb{R}^n \to \mathbb{R}$ Chain Rule)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ and $\underline{G} : \mathbb{R} \to \mathbb{R}^n$, then

$f \circ \underline{G} : \mathbb{R} \to \mathbb{R}$

$$
\begin{aligned}
(f \circ \underline{G})'(t) &= f'(\underline{G}(t))\underline{G}'(t) \\
&= (\nabla f)(\underline{G}(t)) \cdot \underline{G}'(t) \\
&= \sum_{i=1}^{n} \underbrace{\left.\frac{\partial f}{\partial x_i}\right|_{\underline{G}(t)} \left.\frac{dG_i}{dt}\right|_t}.
\end{aligned}
$$

Proof. ∎

Omit.

$$\underbrace{\frac{\partial f}{\partial x_i}(\underline{G}(t))} \cdot \underbrace{\frac{dG_i}{dt}}$$

$$\nabla f(\underline{G}(t)) \cdot \frac{dG}{dt}$$

Example. Consider $f(x, y, z) = xy + z$ and $\underline{G}(t) = \begin{pmatrix} t \\ 1-t \\ e^t \end{pmatrix}$. Find $\cancel{g'(t)}$. $(f \circ \underline{G})'(t)$

$f : \mathbb{R}^3 \longrightarrow \mathbb{R}$

$\underline{G} : \mathbb{R} \rightarrow \mathbb{R}^3$

$f \circ \underline{G} : \mathbb{R} \longrightarrow \mathbb{R}$

$$(f \circ \underline{G})'(t) = \nabla f(\underline{G}(t)) \cdot \underline{G}'(t) = \left.\begin{pmatrix} y \\ x \\ 1 \end{pmatrix}\right|_{\substack{x \to t \\ y \to 1-t \\ z \to e^t}} \cdot \begin{pmatrix} 1 \\ -1 \\ e^t \end{pmatrix}$$

$$= \begin{pmatrix} 1-t \\ t \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ e^t \end{pmatrix} = 1 - t - t + e^t = 1 - 2t + e^t$$

$$(f \circ \underline{G})(t) = f(t, \underline{1-t}, e^t) = t - t^2 + e^t$$

$$(f \circ \underline{G})'(t) = 1 - 2t + e^t \ \checkmark$$

Example. Consider $f : \mathbb{R}^3 \to \mathbb{R}$ and $g(t) = f(t, x(t), y(t))$. Find $g'(t)$.

$x, y : \mathbb{R} \to \mathbb{R}$

$$g(t) = f\begin{pmatrix} t \\ x(t) \\ y(t) \end{pmatrix} = (f \circ \underline{H})(t) \quad \text{where} \quad \underline{H}(t) = \begin{pmatrix} t \\ x(t) \\ y(t) \end{pmatrix}.$$

$f(t, x, y)$.

$$g'(t) = \nabla f(\underline{H}(t)) \cdot \underline{H}'(t)$$

$$= \begin{pmatrix} \frac{\partial f}{\partial t}(H(t)) \\ \frac{\partial f}{\partial x}(\underline{H}(t)) \\ \frac{\partial f}{\partial y}(\underline{H}(t)) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ x'(t) \\ y'(t) \end{pmatrix}$$

$$= \frac{\partial f}{\partial t}(t, x(t), y(t)) + \frac{\partial f}{\partial x}(t, x(t), y(t)) \cdot x'(t) + \frac{\partial f}{\partial y}(t, x(t), y(t)) \cdot y'(t).$$

$$= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \quad \text{in short-hand.}$$

Multivariable Calculus

MATH2007

1.3 The Chain Rule
(Part 2)

WITS UNIVERSITY

School of Mathematics

Example. Given that $f : \mathbb{R}^2 \to \mathbb{R}$ where $f = f(y_1, y_2)$ is constant on the curves parametrized by $\underline{r}_a(t) = \begin{pmatrix} t^2 \\ a + e^t \end{pmatrix}$ for all $a \in \mathbb{R}$, give a differential equation (not unique) for f in terms of y_1, y_2.

fixed.

for all t

$$f(\underline{r}_a(t)) = c \qquad c \text{ is a constant}$$

$$\frac{d}{dt} f(\underline{r}_a(t)) = \frac{d}{dt} c = 0$$

$$\nabla f(\underline{r}_a(t)) \cdot \underline{r}_a'(t) = 0 \quad \Longrightarrow \quad \begin{pmatrix} \frac{\partial f}{\partial y_1}(t^2, a+e^t) \\ \frac{\partial f}{\partial y_2}(t^2, a+e^t) \end{pmatrix} \cdot \begin{pmatrix} 2t \\ e^t \end{pmatrix} = 0$$

$$\Rightarrow \quad \frac{\partial f}{\partial y_1}(t^2, a+e^t) \cdot 2t + \frac{\partial f}{\partial y_2}(t^2, a+e^t) \cdot e^t = 0.$$

$$\Rightarrow \quad \frac{\partial f}{\partial y_2}(t^2, a+e^t) = -2te^{-t} \frac{\partial f}{\partial y_1}(t^2, a+e^t).$$

$$\frac{\partial f}{\partial y_2}(t^2, a+e^t) = -2te^{-t}\,\frac{\partial f}{\partial y_1}(t^2, a+e^t).$$

If we identify $y_1 = t^2$ $y_2 = a+e^t$, then

$$\frac{\partial f}{\partial y_2} = -\frac{2\sqrt{y_1}}{y_2 - a}\,\frac{\partial f}{\partial y_2} \qquad \text{or} \qquad 2\sqrt{y_1}\,\frac{\partial f}{\partial y_2} + (y_2 - a)\,\frac{\partial f}{\partial y_1} = 0$$

many other equations can be derived (not all equivalent!)

Example. Show that if f is constant on the curve parametrized by $\underline{r}(t)$, then $\nabla f(\underline{r}(t))$ is orthogonal to the curve at $\underline{r}(t)$.

$$f(\underline{r}(t)) = c \qquad c \text{ is a constant.}$$

$$\Leftrightarrow \frac{d}{dt} f(\underline{r}(t)) = 0$$

$$\Leftrightarrow \underbrace{\nabla f(\underline{r}(t))} \cdot \underbrace{\underline{r}'(t)}_{\text{tangent vector}} = 0$$

Thus ∇f is orthogonal to the curve $\underline{r}(t)$ at all points on the curve.

Example. Given that $g(t) = f(t, x(t, y(t)), y(t))$. Find $g'(t)$ for $f : \mathbb{R}^3 \to \mathbb{R}$.

$x = x(t, y)$

$f(t, x, y)$

$g : \mathbb{R} \to \mathbb{R} \qquad x : \mathbb{R}^2 \to \mathbb{R} \qquad y : \mathbb{R} \to \mathbb{R}$

Let $\underline{H}(t) = \begin{pmatrix} t \\ x(t, y(t)) \\ y(t) \end{pmatrix} \qquad \underline{K}(t) = \begin{pmatrix} t \\ y(t) \end{pmatrix}$

$$\underline{K}'(t) = \begin{pmatrix} 1 \\ y'(t) \end{pmatrix} \qquad \underline{H}'(t) = \begin{pmatrix} 1 \\ (x \circ \underline{k})'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} 1 \\ \nabla x(\underline{k}(t)) \cdot \underline{k}'(t) \\ y'(t) \end{pmatrix}$$

$$g(t) = (f \circ \underline{H})(t) \qquad g'(t) = \nabla f(\underline{H}(t)) \cdot \underline{H}'(t)$$

$$g'(t) = \begin{pmatrix} \frac{\partial f}{\partial t}(t, x(t, y(t)), y(t)) \\ \frac{\partial f}{\partial x}(t, x(t, y(t)), y(t)) \\ \frac{\partial f}{\partial y}(t, x(t, y(t)), y(t)) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \left(\frac{\partial x}{\partial t}(t, y(t)), \frac{\partial x}{\partial y}(t, y(t))\right) \cdot (1, y'(t)) \\ y'(t) \end{pmatrix}$$

$$g'(t) = \begin{pmatrix} \frac{\partial f}{\partial t}(t, x(t,y(t)), y(t)) \\ \frac{\partial f}{\partial x}(t, x(t,y(t)), y(t)) \\ \frac{\partial f}{\partial y}(t, x(t,y(t)), y(t)) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \left(\frac{\partial x}{\partial t}(t, y(t)), \frac{\partial x}{\partial y}(t, y(t))\right) \cdot (1, y'(t)) \\ y'(t) \end{pmatrix}$$

$$= \frac{\partial f}{\partial t}(t, x(t,y(t)), y(t))$$
$$+ \frac{\partial f}{\partial x}(t, x(t,y(t)), y(t)) \cdot \left(\frac{\partial x}{\partial t}(t, y(t)) + \frac{\partial x}{\partial y}(t, y(t))\, y'(t) \right)$$
$$+ \frac{\partial f}{\partial y}(t, x(t,y(t)), y(t)) \cdot y'(t)$$

$$= \frac{\partial f}{\partial t}(t,x,y) + \frac{\partial f}{\partial x}(t,x,y)\left[\frac{\partial x}{\partial t}(t,y) + \frac{\partial x}{\partial y}(t,y)\, y'(t) \right] + \frac{\partial f}{\partial y}(t,x,y) \cdot y'(t).$$

$$g(t) = f(t, x(t,y), y)$$

$$y(t) \qquad x(t, y(t))$$

$$f(t, x, y)$$

$$g'(t) = \frac{\partial f}{\partial t}\Big|_{\dots} \cdot 1 + \frac{\partial f}{\partial x}\Big|_{\dots} x'(t, y(t)) + \frac{\partial f}{\partial y}\Big|_{\dots} \cdot y'(t)$$

$$= \frac{\partial f}{\partial t}\Big|_{\dots} + \frac{\partial f}{\partial x}\Big|_{\dots} \left[\frac{\partial x}{\partial t}(t, y(t)) \cdot 1 + \frac{\partial x}{\partial y}(t, y(t)) \cdot y'(t) \right] + \frac{\partial f}{\partial y}\Big|_{\dots} \cdot y'(t)$$

$$= \frac{\partial f}{\partial t}(t, x, y) + \frac{\partial f}{\partial x}(t, x, y) \left[\frac{\partial x}{\partial t}(t, y) + \frac{\partial x}{\partial y}(t, y)\, y'(t) \right] + \frac{\partial f}{\partial y}(t, x, y) \cdot y'(t).$$

Multivariable Calculus

MATH2007

1.3 The Chain Rule
(Part 3)

WITS UNIVERSITY

School of Mathematics

Theorem (1.3.3, General Chain Rule)**.** Let $\underline{F} : \mathbb{R}^q \to \mathbb{R}^m$ and $\underline{G} : \mathbb{R}^p \to \mathbb{R}^q$, then

$\underline{F} \circ \underline{G} : \mathbb{R}^p \to \mathbb{R}^m$

$$\underbrace{(\underline{F} \circ \underline{G})'(\underline{x})}_{m \times p} = \underbrace{\underbrace{\underline{F}'(\underline{G}(\underline{x}))}_{m \times q}\underbrace{\underline{G}'(\underline{x})}_{q \times p}}_{\text{order is important.}}.$$

Proof. See tutorial Q6. $\square$

$$(\underline{F} \circ \underline{G})(\underline{x}) \begin{pmatrix} F_1(\underline{G}(\underline{x})) \\ F_2(\underline{G}(\underline{x})) \\ \vdots \\ F_m(\underline{G}(\underline{x})) \end{pmatrix}$$

$$(\underline{F} \circ \underline{G})'(\underline{x}) = \begin{pmatrix} (F_1 \circ \underline{G})'(\underline{x}) \\ (F_2 \circ \underline{G})'(\underline{x}) \\ \vdots \\ (F_m \circ \underline{G})'(\underline{x}) \end{pmatrix}$$

Example. Let

$$\underline{F}(y_1, y_2) = \begin{pmatrix} e^{y_1} \\ y_1 y_2 \\ y_1 - y_2 \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix}$$

and

$$\underline{G}(x_1, x_2) = \begin{pmatrix} \overbrace{\ln x_1}^{=y_1} \\ \underbrace{x_1 - x_2}_{=y_2} \end{pmatrix} = \begin{pmatrix} G_1 \\ G_2 \end{pmatrix}$$

Find $(\underline{F} \circ \underline{G})'$ by using the chain rule.

$$\underline{F}'(y_1, y_2) = \begin{pmatrix} e^{y_1} & 0 \\ y_2 & y_1 \\ 1 & -1 \end{pmatrix}$$

$$\underline{F}'(\underline{G}(x_1, x_2)) = \begin{pmatrix} x_1 & 0 \\ x_1 - x_2 & \ln x_1 \\ 1 & -1 \end{pmatrix}$$

$$\underline{G}'(x_1, x_2) = \begin{pmatrix} \frac{1}{x_1} & 0 \\ 1 & -1 \end{pmatrix}$$

$$(\underline{F} \circ \underline{G})'(x_1, x_2) = \underline{F}'(\underline{G}(x_1, x_2)) \cdot \underline{G}'(x_1, x_2)$$

$$= \begin{pmatrix} x_1 & 0 \\ x_1 - x_2 & \ln x_1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{x_1} & 0 \\ 1 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 1 - \frac{x_2}{x_1} + \ln x_1 & -\ln x_1 \\ \frac{1}{x_1} - 1 & 1 \end{pmatrix}.$$

Note. (a) If $\underline{F} : \mathbb{R}^n \to \mathbb{R}^n$ then $\det(\underline{F}')$ is called the Jacobian of $\underline{F}$ and is denoted by $\frac{\partial \underline{F}(\underline{x})}{\partial \underline{x}}$. i.e $\frac{\partial \underline{F}(\underline{x})}{\partial \underline{x}} = \det(\underline{F}')$.

$$\frac{\partial \underline{F}(\underline{x})}{\partial \underline{x}}$$

(b) If $\underline{F} : \mathbb{R}^n \to \mathbb{R}^n$ and $\underline{G} : \mathbb{R}^n \to \mathbb{R}^n$ then

$$\begin{aligned}
\frac{\partial(\underline{F} \circ \underline{G})(\underline{x})}{\partial \underline{x}} &= \det(\underline{F} \circ \underline{G})'(\underline{x}) \\
&= \det\left[\underline{F}'(\underline{G}(\underline{x}))\underline{G}'(\underline{x})\right] \qquad \text{by chain rule} \\
&= \det \underline{F}'(\underline{G}(\underline{x})) \det \underline{G}'(\underline{x}) \qquad \text{since } \det AB = \det A \det B \\
&= \left.\frac{\partial \underline{F}(\underline{y})}{\partial \underline{y}}\right|_{\underline{y}=\underline{G}(\underline{x})} \frac{\partial \underline{G}(\underline{x})}{\partial \underline{x}}
\end{aligned}$$

Example. Let $\underline{G}(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} = \begin{pmatrix} G_1 \\ G_2 \end{pmatrix}$ and $\underline{F}(x,y) = \begin{pmatrix} x^2 - y^2 \\ xy \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$. Illustrate the Jacobian chain rule.

$$\frac{\partial \underline{F}}{\partial(x,y)}(x,y) = \det\begin{pmatrix} 2x & -2y \\ y & x \end{pmatrix} = 2(x^2+y^2)$$

$$\frac{\partial \underline{G}}{\partial(r,\theta)}(r,\theta) = \det\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r$$

$$\frac{\partial(\underline{F\circ G})}{\partial(r,\theta)} = \frac{\partial \underline{F}}{\partial(x,y)}(\underline{G}(r,\theta))\,\frac{\partial \underline{G}}{\partial(r,\theta)}$$

$$\frac{\partial(F\circ G)}{\partial(r,\theta)}(r,\theta) = 2(x^2+y^2)\Big|_{(x,y)\to(r\cos\theta,\,r\sin\theta)} \cdot r = 2r^3.$$