Single variable chain rule의 한계

$$y = x + x^2$$

$$\frac{dy}{dx} = 1 + 2x$$

$$y = x + x^2$$

$$u1(x) = x^2$$

$$u2(x, u1) = x + u1$$

$$\frac{\partial u1(x)}{\partial x} = 2x$$

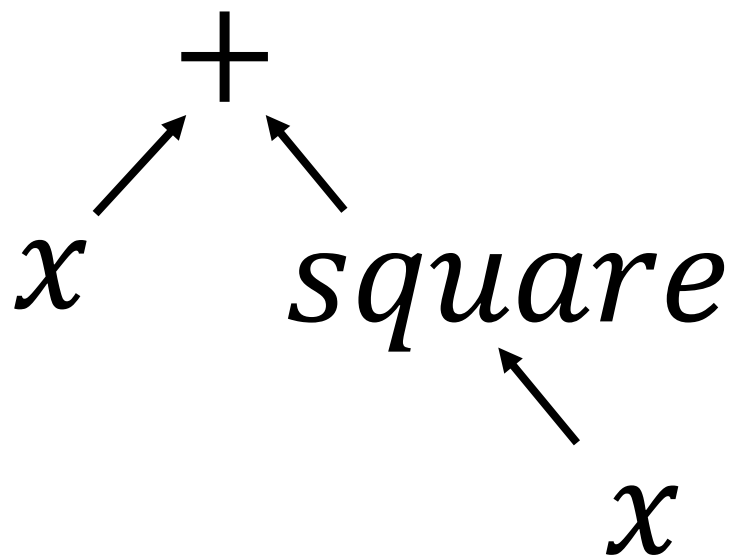$$\frac{\partial u2(x, u1)}{\partial u1} = \frac{\partial}{\partial u1}(x + u1) = 0 + 1 = 1$$

$$\frac{\partial u2(x, u1)}{\partial x} \neq \frac{\partial}{\partial x}(x + u1) = 1 + 0 = 1$$

원인

$$y = x + x^2$$

$$u1(x) = x^2$$

$$u2(x, u1) = x + u1$$

$$+$$

$$y = u2(x, u1)$$

$$x \quad square \quad u1(x)$$

$$x$$

해결방법

single variable total derivative

$$\frac{dy}{dx} = \frac{\partial f(x)}{\partial x} = \frac{\partial u_2(x, u_1)}{\partial x} = \frac{\partial u_2}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial u_2}{\partial u_1}\frac{\partial u_1}{\partial x} = \frac{\partial u_2}{\partial x} + \frac{\partial u_2}{\partial u_1}\frac{\partial u_1}{\partial x}$$

$$\frac{\partial f(x, u_1, \ldots, u_n)}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial u_1}\frac{\partial u_1}{\partial x} + \frac{\partial f}{\partial u_2}\frac{\partial u_2}{\partial x} + \ldots + \frac{\partial f}{\partial u_n}\frac{\partial u_n}{\partial x} = \frac{\partial f}{\partial x} + \sum_{i=1}^{n} \frac{\partial f}{\partial u_i}\frac{\partial u_i}{\partial x}$$

$$y = x + x^2$$

$$u1(x) = x^2$$

$$u2(x, u1) = x + u1$$

$$\frac{dy}{dx} = \frac{\partial f(x)}{\partial x} = \frac{\partial u_2}{\partial x} + \frac{\partial u_2}{\partial u_1}\frac{\partial u_1}{\partial x} = 1 + 1 \times 2x = 1 + 2x$$

예제

$$f(x) = sin(x + x^2)$$

$$
\begin{aligned}
u_1(x) &= x^2 \\
u_2(x, u_1) &= x + u_1 \\
u_3(u_2) &= sin(u_2) \qquad\qquad (y = f(x) = u_3(u_2))
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial u_1}{\partial x} &= 2x \\
\frac{\partial u_2}{\partial x} &= \frac{\partial x}{\partial x} + \frac{\partial u_2}{\partial u_1}\frac{\partial u_1}{\partial x} &= 1 + 1 \times 2x &= 1 + 2x \\
\frac{\partial f(x)}{\partial x} &= \frac{\partial u_3}{\partial x} + \frac{\partial u_3}{\partial u_2}\frac{\partial u_2}{\partial x} &= 0 + cos(u_2)\frac{\partial u_2}{\partial x} &= cos(x + x^2)(1 + 2x)
\end{aligned}
$$

머신러닝에 적용하기 위한 프로세스

$$y = x \times x^2$$

$$u_1(x) = x^2$$
$$u_2(x, u_1) = xu_1 \qquad (y = f(x) = u_2(x, u_1))$$

$$\frac{\partial u_1}{\partial x} = 2x$$
$$\frac{\partial u_2}{\partial x} = u_1 \qquad \left(\text{for } u_2 = x + u_1, \frac{\partial u_2}{\partial x} = 1\right)$$
$$\frac{\partial u_2}{\partial u_1} = x \qquad \left(\text{for } u_2 = x + u_1, \frac{\partial u_2}{\partial u_1} = 1\right)$$

$$\frac{dy}{dx} = \frac{\partial u_2}{\partial x} + \frac{\partial u_2}{\partial u_1}\frac{du_1}{\partial x} = u_1 + x2x = x^2 + 2x^2 = 3x^2$$

용어 사용 시 주의사항