합성함수의 미분법 요약 (연쇄법칙:chain rule)

: 머신러닝을 위한 수학

김종혜

시작하면서...

cost function, 경사하강법, backward를 이해하려면 편미분, 체인룰에 대한 선이해가 필요합니다.

오래간만에 미분을 보시는 분들을 위해 간단히 정리를 해 볼까 합니다.

합성함수

- 두 개 이상의 함수를 하나의 함수로 결합하여 만들어진 함수
- 어떤 함수 속에 또 다른 함수가 들어있고, 그 또 다른 함수 속에 다른 함수가 들어있는 함수

- $f(g(x)) = (f \circ g)(x)$

- 함수 $y = f(u)$ 의 u가 $g(x)$라면, $y = f(g(x))$ 로 중첩해 표현할 수 있음
- 함수 $f(u)$와 $g(x)$의 합성함수

(예) $z = (2-y)^2$ 는 $u = 2 - y$ 함수와 $z = u^2$ 의 합성함수임

$$y \xrightarrow{\quad} u \xrightarrow{\quad} z$$

$$(u = 2 - y) \qquad (z = u^2)$$

합성함수의 미분법(연쇄법칙)

- 합성함수의 미분은 합성함수를 구성하는 각 함수의 미분으로 곱으로 나타냄

(1) 일변수 함수 $x = g(t)$의 변수일 때 $w = f(x)$의 t에 대한 변화

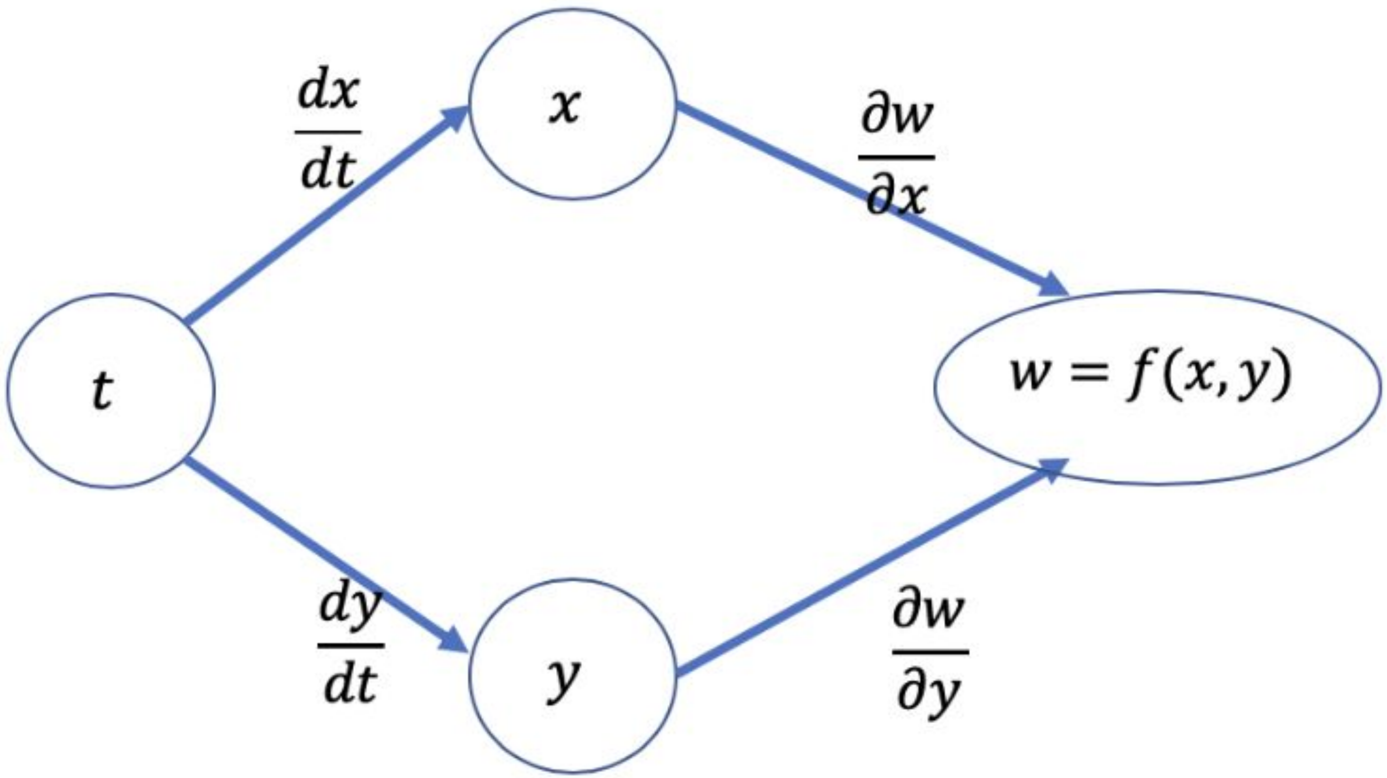
$$\frac{dw}{dt} = \frac{dw}{dx}\frac{dx}{dt}$$

$$t \xrightarrow{\frac{dx}{dt}} x \xrightarrow{\frac{dw}{dx}} w = f(x)$$

t : 독립변수, x : 중간변수, w : 종속변수

(2) $y = f(t), x = g(t)$ 일 때, $w = f(x, y)$의 t에 대한 변화

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}$$



(3) $z = f_1(t), y = f_2(t), x = f_3(t)$일 때, $w = f(x, y, z)$의 t에 대한 변화

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt}$$