CS5242 Neural Networks and Deep Learning

Tutorial 01

Wei WANG

cs5242@comp.nus.edu.sg

NUS National University of Singapore | School of Computing

Gradient for univariate composite functions

- $f(x) = g(x) + h(x)$
  - $f(x) = 3x + x^2$
- $f(x) = \overset{x}{g(x)}\overset{x^2}{h(x)}$
  - $f(x) = xx^2$
- $f(x) = \dfrac{g(x)}{h(x)}$ $\quad e^x$, x
  - $f(x) = \dfrac{e^x}{x}$
- $f(x) = \overset{\ln(u)}{g(u)}, u = \overset{x+x^2}{h(x)}$
  - $f(x) = \ln(x + x^2)$

- $f'(x) = g'(x) + h'(x)$

$f'(x) = 3 + 2x$

- $f'(x) = g'(x)h(x) + g(x)h'(x)$

$f'(x) = 1 \cdot \cancel{2}x^2 + x \cdot 2x = 3x^2$

- $f'(x) = \dfrac{g'(x)h(x) - g(x)h'(x)}{h(x)^2}$

$f'(x) = \dfrac{e^x x - e^x \cdot 1}{x^2} = \dfrac{e^x x - e^x}{x^2}$

- $f'(x) = g'(u)h'(x)$

$f'(x) = \dfrac{1}{x+x^2} \cdot (1+2x) = \dfrac{1+2x}{x+x^2}$

- If $y = (\boldsymbol{Ax})^T\,(2\boldsymbol{x} + \boldsymbol{z})$, where **A** is a square matrix, **x** and **z** are vectors, y is a scalar. what is $\frac{\partial y}{\partial \boldsymbol{x}}$?

$u = Ax$, $v = 2x + z$

$$\frac{\partial y}{\partial x} = \frac{\partial u^T v}{\partial x} = \frac{\partial v}{\partial x} u + \frac{\partial u}{\partial x} \cdot v$$

$$= 2I(Ax) + A^T(2x + z)$$

$$= 2Ax + 2A^T x + A^T z.$$

∂

$$\frac{\partial v}{\partial x} = \frac{\partial\, 2x + z}{\partial x} = 2I$$

$$\frac{\partial u}{\partial x} = \frac{\partial Ax}{\partial x} = A^T$$

- $L = \frac{1}{2}\left(\boldsymbol{w}^T\boldsymbol{x} - y\right)^2$ if $\boldsymbol{x} = (1, 2)^T$, $\boldsymbol{w} = (2,\ 1)^T$, $y = \ 0$. Compute the gradient of $\frac{\partial L}{\partial \boldsymbol{w}}$

$$L = \frac{1}{2} u^T u \quad \text{where} \quad u = w^T x - y$$

$$\frac{\partial L}{\partial w} = \frac{1}{2}\frac{\partial u^T u}{\partial w} = u \cdot \frac{\partial u}{\partial w} = (w^T x - y) \cdot x = 4 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}$$

$$\cancel{(2, 1)^T}$$

$$\frac{\partial u}{\partial w} = \frac{\partial w^T x}{\partial w} = x \qquad (2, 1)\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\underline{u = 4}$$