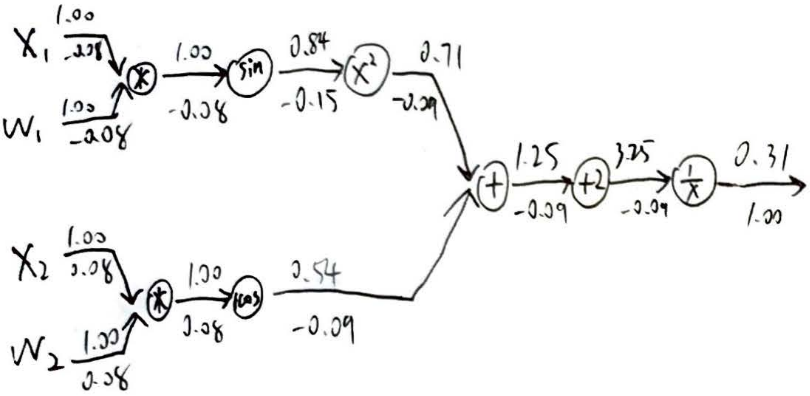
Xin Yang

xy213

HW2:

Problem 1: $f(x,w) = \dfrac{1}{2+\sin^2(x_1 w_1)+\cos(x_2 w_2)}$



With inputs: $x_1=1.00$, $w_1=1.00$, $x_2=1.00$, $w_2=1.00$,

Gradient with respect to x_1: -0.08

w_1: -0.08

x_2: 0.08

w_2: 0.08.

~~$1.00 \cdot \frac{-1}{3.25^2} = -0.09$~~

$f(x)=\frac{1}{x}$ $\quad \frac{df}{dx} = -\frac{1}{x^2}$ $\quad (1.00)\left(\frac{-1}{3.25^2}\right) = -0.09$

$f_c(x)=c+x$ $\quad \frac{df}{dx}=1$ $\quad (-0.09)(1) = -0.09$

$f(x,y)=x+y$ $\quad \frac{df}{dx}=1, \frac{df}{dy}=1$ $\quad (-0.09)(1)=-0.09$

$(-0.09)(1)=-0.09$

$f(x)=x^2$ $\quad \frac{df}{dx}=2x$ $\quad (-0.09)(2\times 0.84) = -0.15$

$f(x)=\sin x$ $\quad \frac{df}{dx}=\cos x$ $\quad (-0.15)(\cos(1.00)) = -0.08$

$f(x,y)=xy$, $\quad \frac{df}{dx}=y, \frac{df}{dy}=x.$ $\quad (-0.08)\cdot 1.00 = -0.08$

$(-0.08)\cdot 1.00 = -0.08$

$f(x)=\cos x$, $\quad \frac{df}{dx}=-\sin x$ $\quad (-0.09)\cdot(-\sin(1.00)) = 0.08$

$f(x,y)=xy$, $\quad \frac{df}{dx}=y, \frac{df}{dy}=x,$ $\quad (0.08)\cdot(1.00)=0.08$

$(0.08)\cdot(1.00)=0.08$

Xin Yang

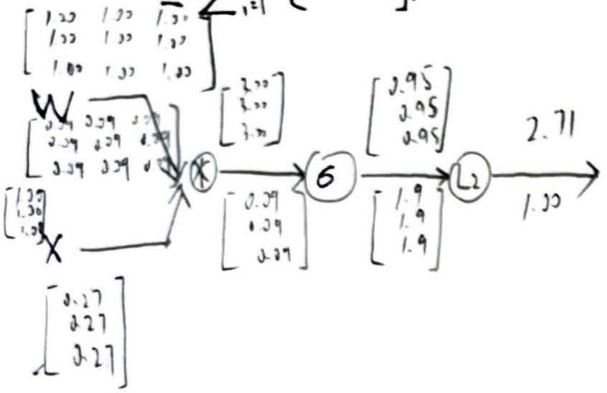
xy213

Hw2.

Problem 2:

$f(x,w) = ||\sigma(W_x)||^2$, $W: 3\times3$, $x: 3\times1$

$$= \sum_{i=1}^{n} [\sigma(W\cdot x)]_i^2$$



$$q = \sigma(Wx) = \begin{pmatrix} \sigma(W_{1,1}x_1) + \cdots + \sigma(W_{1,n}x_n) \\ \vdots \\ \sigma(W_{n,1}x_1) + \cdots + \sigma(W_{n,n}x_n) \end{pmatrix}$$

$$f(q) = ||q||^2 = q_1^2 + q_2^2 + \cdots + q_n^2$$

$$\frac{\partial f}{\partial q_i} = 2q_i \quad \nabla_q f = 2q \quad (1.00)\left(2\begin{bmatrix} 0.95 \\ 0.95 \\ 0.95 \end{bmatrix}\right) = \begin{bmatrix} 1.9 \\ 1.9 \\ 1.9 \end{bmatrix}$$

$m = Wx$

$q = \sigma(m)$

$$\frac{dq}{dm} = \frac{d\sigma(m)}{dm} = \frac{e^{-m}}{(1+e^{-m})^2} = (1-\sigma(m))\,\sigma(m) \quad \left(\begin{bmatrix} 1.9 \\ 1.9 \\ 1.9 \end{bmatrix}\right)\Big((1-\sigma(3.0))(\sigma(3.0))\Big) = \begin{bmatrix} 0.09 \\ 0.09 \\ 0.09 \end{bmatrix}$$

$$\frac{\partial m_k}{\partial W_{ij}} = 1_{k=i}\, x_j \qquad \begin{bmatrix} 0.09 & 0.09 & 0.09 \\ 0.09 & 0.09 & 0.09 \\ 0.09 & 0.09 & 0.09 \end{bmatrix}$$

For inputs: $W = \begin{bmatrix} 1.00 & 1.00 & 1.00 \\ 1.00 & 1.00 & 1.00 \\ 1.00 & 1.00 & 1.00 \end{bmatrix}$

$X = \begin{bmatrix} 1.00 \\ 1.00 \\ 1.00 \end{bmatrix}$

Gradient respect to W: $\begin{bmatrix} 0.09 & 0.09 & 0.09 \\ 0.09 & 0.09 & 0.09 \\ 0.09 & 0.09 & 0.09 \end{bmatrix}$

$$\frac{\partial m_k}{\partial x_i} = W_{ki} \qquad \begin{bmatrix} 0.27 \\ 0.27 \\ 0.27 \end{bmatrix}$$

X: $\begin{bmatrix} 0.27 \\ 0.27 \\ 0.27 \end{bmatrix}$