Deep Learning **HW #4** **Instructor: Shusen Pu**

Student Name (print): Farooq Mahmud **Group #** 10

Question:	1	2	3	4	Total
Points:	10	10	20	0	40
Bonus Points:	0	0	0	10	10
Score:					

1. (10 points) Find the derivatives of $f(x)$ using basic rules of differentiation.

   (a) $f(x) = \sin x \cos x$

$$= [sin(x)(-sin(x))] + [cos(x)cos(x)]$$ (product rule)

$$= -sin^2(x) + cos^2(x)$$

   (b) $f(x) = \tan x$ (hint: $\tan x = \frac{\sin x}{\cos x}$)

$$= \frac{[cos(x)cos(x)] - [sin(x)(-sin(x))]}{cos^2(x)}$$ (quotient rule)

$$= \frac{cos^2(x) + sin^2(x)}{cos^2(x)}$$

$$= \frac{1}{cos^2(x)}$$ (Pythagorean identity)

$$= sec^2(x)$$

(c) $f(x) = \sin^3(x^2)$ (Note that $\sin^3(x) = [\sin x]^3$)

$$= (sin(x^2))^3 \qquad \text{(power rule)}$$
$$= (3sin^2(x^2))(sin'(x^2)) \qquad \text{(chain rule)}$$
$$= (3sin^2(x^2))(2xcos(x^2))$$
$$= 6xsin^2(x^2)cos(x^2)$$

(d) $f(x) = x\sin(x^3)$

$$= x(sin'(x^3)) + (1(sin(x^3))) \qquad \text{(product rule)}$$
$$= x(cos(x^3)3x^2)) + sin(x^3) \qquad \text{(chain rule)}$$
$$= 3x^3cos(x^3) + sin(x^3)$$

2. (10 points) Let $f(x, y, z) = w_1 x^3 + w_2 x \mathrm{e}^y + w_3 x y z^2$, where w_1, w_2 and w_3 are weight constants. Find the following partial derivatives.

   (a) $\frac{\partial f}{\partial x}$

   $$3x^2 w_1 + w_2 e^y + w_3 y z^2$$

   (b) $\frac{\partial f}{\partial y}$

   $$w^2 x e^y + w_3 x z^2$$

   (c) $\frac{\partial f}{\partial z}$

   $$2z w_3 x y$$

   (d) Find the gradient of $f(x, y, z)$, i.e., $\nabla f(x, y, z)$

   $$3x^2 w_1 + w_2 e^y + w_3 y z^2 + w^2 x e^y + w_3 x z^2 + 2z w_3 x y$$

   (e) What is the gradient of $f(x, y, z)$ at $(1, 2, 1)$

$$3(1^2)w_1 + w_2 e^2 + w_3(2)(1^2) + w^2(1)e^2 + w_3(1)(1^2) + 2(1)w_3(1)(2)$$
$$= 3w_1 + w_2 e^2 + 2w_3 + w^2 e^2 + w_3 + 4w_3$$

3. (20 points) Suppose a 3×3 matrix W has an eigenvalue decomposition as follows:

$$W = X \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} X^{-1} \tag{1}$$

where X is the square $n \times n$ matrix whose ith column is the eigenvector q_i of W, λ_1, λ_2 and λ_3 are eigenvalues of W arranged in descending order.

F.Y.I: matrix multiplication is defined in the following way:

> If $A = (a_{ij})$ is an $m \times n$ matrix and $B = (b_{ij})$ is an $n \times r$ matrix, then the product $AB = C = (c_{ij})$ is the $m \times r$ matrix whose entries are defined by
>
> $$c_{ij} = \vec{\mathbf{a}}_i \mathbf{b}_j = \sum_{k=1}^{n} a_{ik} b_{kj}$$

(a) Find W^3 use equation(1) and the definition of matrix product. Show your steps.

$$W^3 = W \times W \times W$$

$$= \left[X \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} X^{-1}\right]\left[X \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} X^{-1}\right]\left[X \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} X^{-1}\right]$$

$$= X \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \cancel{(XX^{-1})} X^{-1} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

$$X \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} X^{-1}$$

$$= X \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \cancel{X^{-1}X} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} X^{-1}$$

$$= X \begin{pmatrix} \lambda_1^3 & 0 & 0 \\ 0 & \lambda_2^3 & 0 \\ 0 & 0 & \lambda_3^3 \end{pmatrix} X^{-1}$$

(b) Let $\lambda_1 = 3, \lambda_2 = 2$ and $\lambda_3 = 1$, find W^{30}

If W^3 raises each λ into the 3rd power,
W^{30} raises each λ to the 30th power

$$W^{30} = X \begin{pmatrix} 3^{30} & 0 & 0 \\ 0 & 2^{30} & 0 \\ 0 & 0 & 1^{30} \end{pmatrix} X^{-1}$$

$$= X \begin{pmatrix} 2.06 \times 10^{14} & 0 & 0 \\ 0 & 1.07 \times 10^{9} & 0 \\ 0 & 0 & 1 \end{pmatrix} X^{-1}$$

(c) Let $\lambda_1 = 0.3, \lambda_2 = 0.2$ and $\lambda_3 = 0.1$, find W^{30}

$$W^{30} = X \begin{pmatrix} 2.06 \times 10^{-16} & 0 & 0 \\ 0 & 1.07 \times 10^{-21} & 0 \\ 0 & 0 & 1 \end{pmatrix} X^{-1}$$

(d) Note: this is an example of how the long-time dependence in RNN can lead to exploding or vanishing gradient.

4. (10 points (bonus)) Use the definition $f'(x) = \lim_{h \to 0} \dfrac{f(x+h) - f(x)}{h}$ to find $f'(x)$ for the following functions.

   a. $f(x) = x^2$
   b. $f(x) = x^3$
   c. $f(x) = x^4$
   d. Based on your answers to parts a-c, propose a formula for $f'(x)$ if $f(x) = x^n$, where n is a positive integer.

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1
\end{array}
$$

(a)

$$f(x) = x^2$$
$$= \frac{(x+h)^2 - x^2}{h} = \frac{x^2 + 2xh + h^2 - x^2}{h} = \frac{h^2 + 2xh}{h} = \frac{h(h+2x)}{h} = h + 2x$$
$$\lim_{h \to 0}(h + 2x) = 2x$$

(b)

$$f(x) = x^3$$
$$= \frac{(x+h)^3 - x^3}{h} = \frac{x^3 + 3x^2h + 3xh^2 + h^3 - x^3}{h} = \frac{h(3x^2 + 3xh + h^2)}{h} = 3x^2 + 3xh + h^2$$
$$\lim_{h \to 0}(3x^2 + 3xh + h^2) = 3x^2$$

(c)

$$f(x) = x^4$$
$$= \frac{(x+h)^4 - x^4}{h} = \frac{x^4 + 4x^3h + 6x^2h^2 + 4xh^3 + h^4 - x^4}{h} - \frac{4x^3h + 6x^2h^2 + 4xh^3 + h^4}{h}$$
$$= \frac{h(4x^3 + 6x^2h + 4xh^2 + h^3)}{h}$$
$$= 4x^3 + 6x^2h + 4xh^2 + h^3$$
$$\lim_{h \to 0}(4x^3 + 6x^2h + 4xh^2 + h^3) = 4x^3$$

(d) $f(x) = nx^{n-1}$