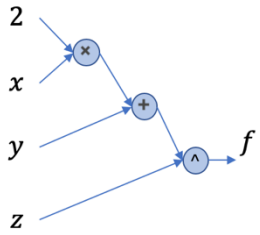
Name: _____Student ID:_____

11040A Neural Networks

Assignment 3

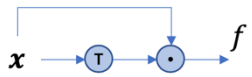
Deadline: May 29, 2025 24:00:00

1. Composite functions can be depicted as computational graphs.
   For example, the figure below shows the corresponding computational graph of function $f(x, y, z) = z^{2x+y}$.



   Draw the computational graphs for the following functions and calculate the partial derivatives of all the functions with respect to x_1.
   (a) $f(x_1, x_2) = (1 + e^{w_1x_1+w_2x_2+b})^{-1}$
   (b) $f(x_1, x_2) = \frac{e^{x_1}}{e^{x_1}+e^{x_2}}$
   (c) $f(x_1, x_2) = \ln\frac{e^{x_1}}{e^{x_1}+e^{x_2}}$
   (d) $f(x_1, x_2) = \ln\frac{e^{x_2}}{e^{x_1}+e^{x_2}}$
   (e) $f(x_1, x_2, x_3, x_4) = \ln\frac{e^{w_1x_1+w_2x_2}}{e^{w_1x_1+w_2x_2}+e^{w_1x_3+w_2x_4}}$
   (f) $f(x_1, x_2, x_3, x_4) = \ln\frac{e^{w_1x_3+w_2x_4}}{e^{w_1x_1+w_2x_2}+e^{w_1x_3+w_2x_4}}$

2. For simplicity, a node can also represent a vector or a matrix in the computational graph.
   For example, function $f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{x}$, where $\boldsymbol{x} = [x_1, x_2]^T$ can be described as the figure below.



   Draw the computational graphs for the following functions and calculate the gradients of all the functions with respect to $\boldsymbol{x}$.
   (a) $f(\boldsymbol{x}) = (\boldsymbol{\theta}^T\boldsymbol{x})^T\boldsymbol{\theta}^T\boldsymbol{x}$, where $\boldsymbol{\theta} \in \mathbb{R}^{d\times h}$ and $\boldsymbol{x} \in \mathbb{R}^{d\times 1}$.
   (b) $f(\boldsymbol{x}) = \boldsymbol{\theta}^T(\boldsymbol{x} + \boldsymbol{\theta}^T\boldsymbol{x})$, where $\boldsymbol{\theta} \in \mathbb{R}^{d\times d}$ and $\boldsymbol{x} \in \mathbb{R}^{d\times 1}$.
   (c) $f(\boldsymbol{x}) = (\boldsymbol{a}^T\boldsymbol{x})\boldsymbol{a} + (\boldsymbol{b}^T\boldsymbol{x})\boldsymbol{b}$, where $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{x} \in \mathbb{R}^{d\times 1}$.

   Notice that matrix calculus has largely two consistent layout conventions: *Numerator layout* and *Denominator layout*. Therefore, sometimes, the question may have more than one answer. (You may refer to the "Matrix calculus" entry on Wikipedia for more information.)