Instructions:

- This assignment is meant to help you grok certain concepts we will use in the course. Please don't copy solutions from any sources.

- Avoid verbosity.

- The assignment needs to be written in latex using the attached tex file. The solution for each question should be written in the solution block in space already provided in the tex file. **Handwritten assignments will not be accepted.**

1. **Partial Derivatives**

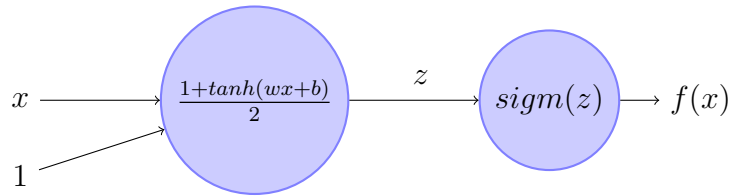   (a) Find the derivative of $g(\rho)$ with respect to ρ where $g(\rho)$ is given by,

   $$g(\rho) = \frac{1}{2}\rho log \frac{\rho}{\rho + \hat{\rho}} + \frac{1}{2}\hat{\rho} log \frac{\hat{\rho}}{\rho + \hat{\rho}}$$

   (You can consider $\hat{\rho}$ as constant)

   > **Solution:** The derivative of $g(\rho)$ with respect to ρ can be found as follows:
   >
   > $$g'(\rho) = \frac{d}{d\rho}(\frac{1}{2}(\rho log\rho - \rho log(\rho + \hat{\rho})) + \frac{1}{2}\hat{\rho}(log\hat{\rho} - log(\rho + \hat{\rho})))$$
   >
   > $$= \frac{1}{2}(\frac{d\rho}{d\rho}log\rho + \rho\frac{d}{d\rho}(log\rho) - \rho\frac{d}{d\rho}(log(\rho + \hat{\rho})) - \frac{d\rho}{d\rho}log(\rho + \hat{\rho})) - \frac{1}{2}\hat{\rho}\frac{d}{d\rho}(log(\rho + \hat{\rho}))$$
   >
   > $$(By\ Product\ Rule)$$
   >
   > $$= \frac{1}{2}(log\rho + \frac{\rho}{\rho} - \frac{\rho}{\rho + \hat{\rho}} - log(\rho + \hat{\rho}) - \frac{\hat{\rho}}{\rho + \hat{\rho}})$$
   >
   > $$= \frac{1}{2}log(\frac{\rho}{\rho + \hat{\rho}})$$

   (b) Consider the following computation ,

   where $z = \frac{1+tanh(wx+b)}{2}$ and $f(x) = sigm(z)$

   by definition : $sigm(z) = \frac{1}{1+e^{-z}}$ and $tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$

   The value L is given by,
   $$L = -y \log(f(x))$$

Here, x and y are constants and w and b are parameters that can be modified. In other words, L is a function of w and b.

Derive the partial derivatives, $\frac{\partial L}{\partial w}$ and $\frac{\partial L}{\partial b}$.

Solution: Given, $L = -y \log(f(x)) = -y \log(sigm(z))$

The partial derivative of L w.r.t w is given by $\frac{\partial L}{\partial w}$, which can be found as follows:

$$\frac{\partial L}{\partial w} = -y \frac{\partial}{\partial w}(log(sigm(z)))$$

$$= -y \frac{1}{sigm(z)} \frac{\partial}{\partial w}(sigm(z)) \qquad (By\ Chain\ Rule)$$

The derivative of the sigmoid function is : sigm'(z) = sigm(z)(1-sigm(z))

$$\frac{\partial L}{\partial w} = -y \frac{1}{sigm(z)}(sigm(z)(1 - sigm(z)))\frac{\partial z}{\partial w} \qquad (By\ Chain\ Rule)$$

$$= -y(1 - sigm(z))(\frac{1}{2}\frac{\partial}{\partial w}tanh(wx + b))$$

$$= -xy(1 - sigm(z))(\frac{1}{2}sech^2(wx + b))$$

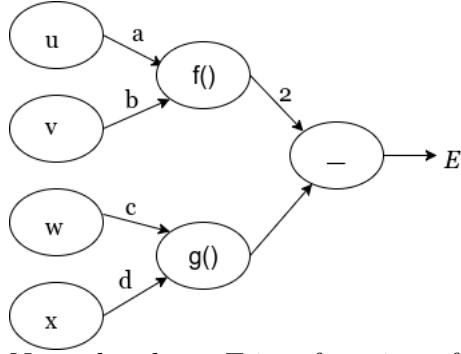The partial derivative of L w.r.t b is given by $\frac{\partial L}{\partial b}$, which can be found as follows:

$$\frac{\partial L}{\partial b} = -y \frac{1}{sigm(z)}(sigm(z)(1 - sigm(z)))\frac{\partial z}{\partial b} \qquad (By\ Chain\ Rule)$$

$$= -y(1 - sigm(z))(\frac{1}{2}\frac{\partial}{\partial b}tanh(wx + b))$$

$$= -y(1 - sigm(z))(\frac{1}{2}sech^2(wx + b))$$

---

2. **Chain Rule**:

   (a) Consider the evaluation of E as given below,

$$E = h(u, v, w, x) = 2 * f(au + bv) - g(cw + dx)$$

Represented as graph:

Here u, v, w, x are inputs (constants) and a, b, c, d are parameters (variables). f and g are the activation functions (with z as input) defined as below:

$$f(z) = sigm(z) \qquad\qquad g(z) = tanh(z)$$

Note that here E is a function of parameters a, b, c, d. Compute the partial derivatives of E with respect to the parameters a, b, c and d i.e. $\frac{\partial E}{\partial a}$, $\frac{\partial E}{\partial b}$, $\frac{\partial E}{\partial c}$ and $\frac{\partial E}{\partial d}$.

Solution: Given, E = 2*f(au + bv) - g(cw + dx) , f(z) = sigm(z) and g(z) = tanh(z).

$$\therefore E = 2 * sigm(au + bv) - tanh(cw + dx)$$

The partial derivative of E w.r.t a is given by $\frac{\partial E}{\partial a}$, which can be found as follows:

$$\frac{\partial E}{\partial a} = 2 * \frac{\partial}{\partial a}(sigm(au + bv))$$

$$= 2 * sigm(au + bv)(1 - sigm(au + bv))\frac{\partial}{\partial a}(au + bv) \qquad (By\,Chain\,Rule)$$

$$= 2u * sigm(au + bv)(1 - sigm(au + bv))$$

The partial derivative of E w.r.t b is given by $\frac{\partial E}{\partial b}$, which can be found as follows:

$$\frac{\partial E}{\partial b} = 2 * \frac{\partial}{\partial b}(sigm(au + bv))$$

$$= 2 * sigm(au + bv)(1 - sigm(au + bv))\frac{\partial}{\partial b}(au + bv) \qquad (By\,Chain\,Rule)$$

$$= 2v * sigm(au + bv)(1 - sigm(au + bv))$$

The partial derivative of E w.r.t c is given by $\frac{\partial E}{\partial c}$, which can be found as follows:

$$\frac{\partial E}{\partial c} = -\frac{\partial}{\partial c}(tanh(cw + dx))$$

$$= -sech^2(cw + dx)\frac{\partial}{\partial c}(cw + dx) \qquad (By\,Chain\,Rule)$$

$$= -w * sech^2(cw + dx)$$

The partial derivative of E w.r.t d is given by $\frac{\partial E}{\partial d}$, which can be found as follows:

$$\frac{\partial E}{\partial d} = -\frac{\partial}{\partial d}(tanh(cw + dx))$$

$$= -\text{sech}^2(cw + dx)\frac{\partial}{\partial d}(cw + dx) \qquad\qquad (By\,Chain\,Rule)$$

$$= -x * \text{sech}^2(cw + dx)$$

(b) Assume that $z = f(x, y)$, where $x = uv$ and $y = \frac{u}{v}$. We use f_x to denote the partial derivative $\frac{\partial f}{\partial x}$. Using the chain rule, express $\frac{\partial z}{\partial u}$ and $\frac{\partial z}{\partial v}$ in terms of u, v, f_x and f_y.

Solution: Given, $z = f(x, y)$

The partial derivative of z w.r.t u is given by $\frac{\partial z}{\partial u}$, which can be found as follows:

$$\frac{\partial z}{\partial u} = \frac{\partial f}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} \qquad\qquad (By\,Chain\,Rule)$$

Substituting $\frac{\partial f}{\partial x}$ with f_x , $\frac{\partial f}{\partial y}$ with f_y , x with uv and y with $\frac{u}{v}$,

$$= f_x\frac{\partial}{\partial u}(uv) + f_y\frac{\partial}{\partial u}(\frac{u}{v})$$

$$= f_x v + f_y\frac{1}{v}$$

The partial derivative of z w.r.t v is given by $\frac{\partial z}{\partial v}$, which can be found as follows:

$$\frac{\partial z}{\partial v} = \frac{\partial f}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v} \qquad\qquad (By\,Chain\,Rule)$$

$$= f_x\frac{\partial}{\partial v}(uv) + f_y\frac{\partial}{\partial v}(\frac{u}{v})$$

$$= f_x u - f_y\frac{u}{v^2}$$

(c) Given the change of variables as mentioned in the previous part: $x = uv$ and $y = \frac{u}{v}$, calculate the Jacobian of this transformation.

Solution: The jacobian of the transformation $x = uv$ and $y = \frac{u}{v}$ is:

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

$$= \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v}$$

$$= \frac{\partial}{\partial u}(uv)\frac{\partial}{\partial v}(\frac{u}{v}) - \frac{\partial}{\partial u}(\frac{u}{v})\frac{\partial}{\partial v}(uv)$$

$$= (v)\left(-\frac{u}{v^2}\right) - \left(\frac{1}{v}\right)(u)$$
$$= -2 * \frac{u}{v}$$

(d) Calculate the Jacobian of the transformation for rectangular coordinates; *i.e.*, the Jacobian of $x = r\sin\theta$, $y = r\cos\theta$, $z = z$, (hint: using the relevant partial derivatives)

Solution: The jacobian of the transformation $x = r\sin\theta$, $y = r\cos\theta$ and $z = z$ is:

$$\frac{\partial(x, y, z)}{\partial(r, \theta, z)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial z} \end{vmatrix}$$
$$= \begin{vmatrix} \sin\theta & r\cos\theta & 0 \\ \cos\theta & -r\sin\theta & 0 \\ 0 & 0 & 1 \end{vmatrix}$$
$$= -r * \sin^2\theta - r * \cos^2\theta$$
$$= -r(\sin^2\theta + \cos^2\theta)$$
$$= -r$$

3. **Visit Taylor Series** The first order derivative of a function f is defined by the following limit,

$$\frac{df(x)}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{1}$$

On observing the above definition we see that the derivative of a function is the ratio of change in the function value to the change in the function input, when we change the input by a small quantity (infinitesimally small). A first degree approximation based on eq. 1 would be the following.

$$f(x+h) \approx f(x) + h\frac{df(x)}{dx} \tag{2}$$

Consider $f(x) = ln(x + 5)$.

(a) Estimate the value of $f(1), f(1.1)$ and $f(2.5)$ using the above formula.

Solution: Given,

$$f(x) = ln(x + 5)$$

$$\therefore \frac{df(x)}{dx} = f'(x) = \frac{1}{x+5}$$

First degree approximation is: $f(x+h) = f(x) + hf'(x)$

Let, $x = 0$ and $h = 1$ then,

$$f(1) = f(0) + 1 * f'(0)$$
$$= ln(5) + \frac{1}{5} = 1.8094$$

Let, $x = 0$ and $h = 1.1$ then,

$$f(1.1) = f(0) + 1.1 * f'(0)$$
$$= ln(5) + \frac{1.1}{5} = 1.8294$$

Let, $x = 0$ and $h = 2.5$ then,

$$f(2.5) = f(0) + 2.5 * f'(0)$$
$$= ln(5) + \frac{2.5}{5} = 2.1094$$

(b) Compare these estimates to the actual values of function $f(1), f(1.1)$ and $f(2.5)$. Explain the discrepancy as we increase the value.

Solution: Actual vs estimated values of the function $f(x) = ln(x+5)$:

f	Actual	Estimated
f(1)	1.79175	1.8094
f(1.1)	1.80828	1.8294
f(2.5)	2.01490	2.1094

Here, the values are estimated by taking the derivatives at 0. It gives the tangent line to function at x = 0. So near to $x = 0$ it'll give better accuracy and As we increase the value, distance between tangent line of function and actual value is increases, accuracy is decreases.

(c) Can we get a better estimate of $f(1), f(1.1)$ and $f(2.5)$? How?

Solution: Yes, we can get better estimation of function values by taking the second order approximation as it'll add one more term, which is related to the second derivative.

second order approximation is: $f(x + h) = f(x) + hf'(x) + \frac{h^2}{2} f''(x)$.

$$f(x) = ln(x + 5)$$
$$\therefore \frac{df(x)}{dx} = f'(x) = \frac{1}{x + 5}$$
$$\therefore \frac{d^2 f(x)}{dx^2} = f''(x) = -\frac{1}{(x + 5)^2}$$

Let, $x = 0$ and $h = 1$ then,

$$f(1) = f(0) + 1 * f'(0) + \frac{1}{2} * f''(0)$$
$$= ln(5) + \frac{1}{5} - \frac{1}{50} = 1.7894$$

Let, $x = 0$ and $h = 1.1$ then,

$$f(1.1) = f(0) + 1.1 * f'(0) + \frac{(1.1)^2}{2} f''(0)$$
$$= ln(5) + \frac{1.1}{5} - \frac{1.21}{50} = 1.8052$$

Let, $x = 0$ and $h = 2.5$ then,

$$f(2.5) = f(0) + 2.5 * f'(0) + \frac{(2.5)^2}{2} f''(0)$$
$$= ln(5) + \frac{2.5}{5} - \frac{6.25}{50} = 1.9844$$

Second order approximation gives better estimation compared to first order approximation.

(d) Consider $g(x) = a + be^x + c * \cos(x)$. Find a, b, c $\in$ R such that g approximates f at $x = 0$. (i.e. by matching the (i) direct values (ii) first derivative and (iii) second derivative at $x = 0$).

Solution:

$$g(x) = a + be^x + c * \cos(x)$$
$$\therefore \frac{dg(x)}{dx} = g'(x) = be^x - c * \sin(x)$$
$$\therefore \frac{d^2 g(x)}{dx^2} = g''(x) = be^x - c * \cos(x)$$

And,

$$f(x) = ln(x + 5)$$

$$\therefore \frac{df(x)}{dx} = f'(x) = \frac{1}{x + 5}$$

$$\therefore \frac{d^2 f(x)}{dx^2} = f"(x) = -\frac{1}{(x + 5)^2}$$

By comparing direct value of functions at $x = 0$,

$$g(0) = f(0)$$

$$a + b + c = ln(5) \qquad\qquad\qquad (1)$$

By comparing first derivative of functions at $x = 0$,

$$g'(0) = f'(0)$$

$$b = \frac{1}{5} = 0.2$$

By comparing second derivative of functions at $x = 0$,

$$g"(0) = f"(0)$$

$$b - c = -\frac{1}{25}$$

$$\therefore c = b + \frac{1}{25}$$

$$\therefore c = \frac{6}{25} = 0.24$$

Now, putting values of b and c into equation(1),

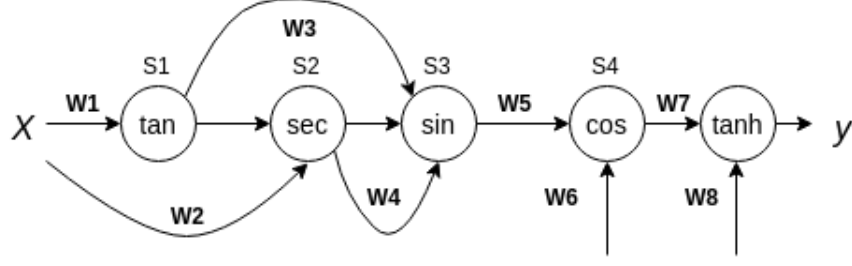$$a = ln(5) - 0.2 - 0.24$$

$$\therefore a = 1.16944$$

4. **Differentiation of function of multiple variables**

$$s_1 = tan(w_1 x)$$

$$s_2 = sec(w_2 x + s_1)$$

$$s_3 = sin(w_3 s_1 + w_4 s_2)$$

$$s_4 = cos(w_5 s_3 + w_6)$$

$$y = tanh(w_7 s_4 + w_8)$$

An alternative representation of the function y is given in the figure below.

Compute the derivatives $\frac{dy}{dw_1}$ and $\frac{dy}{dw_2}$ (show all the steps).

Solution: The derivative of y w.r.t w_1 is given by $\frac{dy}{dw_1}$, which can be found as follows:

$$\frac{dy}{dw_1} = \frac{d}{dw_1}(tanh(w_7 s_4 + w_8))$$

$$= sech^2(w_7 s_4 + w_8)\frac{d}{dw_1}(w_7 s_4 + w_8)$$

$$= sech^2(w_7 s_4 + w_8)(w_7 \frac{d}{dw_1}(cos(w_5 s_3 + w_6)))$$

$$= -w_7 sech^2(w_7 s_4 + w_8)sin(w_5 s_3 + w_6)\frac{d}{dw_1}(w_5 s_3 + w_6)$$

$$= -w_7 sech^2(w_7 s_4 + w_8)sin(w_5 s_3 + w_6)(w_5 \frac{d}{dw_1}sin(w_3 s_1 + w_4 s_2))$$

$$= -w_5 w_7 sech^2(w_7 s_4 + w_8)sin(w_5 s_3 + w_6)cos(w_3 s_1 + w_4 s_2)\frac{d}{dw_1}(w_3 s_1 + w_4 s_2)$$

$$= -w_5 w_7 sech^2(w_7 s_4 + w_8)sin(w_5 s_3 + w_6)cos(w_3 s_1 + w_4 s_2)$$
$$(w_3 \frac{d}{dw_1}tan(w_1 x) + w_4 \frac{d}{dw_1}sec(w_2 x + s_1))$$

$$= -w_5 w_7 sech^2(w_7 s_4 + w_8)sin(w_5 s_3 + w_6)cos(w_3 s_1 + w_4 s_2)$$
$$(w_3 sec^2(w_1 x)\frac{d}{dw_1}(w_1 x) + w_4 sec(w_2 x + s_1)tan(w_2 x + s_1)\frac{d}{dw_1}(w_2 x + s_1))$$

$$= -w_5 w_7 sech^2(w_7 s_4 + w_8)sin(w_5 s_3 + w_6)cos(w_3 s_1 + w_4 s_2)$$
$$(w_3 x sec^2(w_1 x) + w_4 sec(w_2 x + s_1)tan(w_2 x + s_1)\frac{d}{dw_1}(tan(w_1 x)))$$

$$= -w_5 w_7 sech^2(w_7 s_4 + w_8)sin(w_5 s_3 + w_6)cos(w_3 s_1 + w_4 s_2)$$
$$(w_3 x sec^2(w_1 x) + w_4 x sec(w_2 x + s_1)tan(w_2 x + s_1)sec^2(w_1 x))$$

$$= -x w_5 w_7 sech^2(w_7 s_4 + w_8)sin(w_5 s_3 + w_6)cos(w_3 s_1 + w_4 s_2)sec^2(w_1 x)$$
$$(w_3 + w_4 sec(w_2 x + s_1)tan(w_2 x + s_1))$$

The derivative of y w.r.t w_2 is given by $\frac{dy}{dw_2}$, which can be found as follows:

$$\frac{dy}{dw_2} = \frac{d}{dw_2}(tanh(w_7 s_4 + w_8))$$

$$= sech^2(w_7s_4 + w_8)\frac{d}{dw_2}(w_7s_4 + w_8)$$

$$= sech^2(w_7s_4 + w_8)(w_7\frac{d}{dw_2}(cos(w_5s_3 + w_6)))$$

$$= -w_7sech^2(w_7s_4 + w_8)sin(w_5s_3 + w_6)\frac{d}{dw_2}(w_5s_3 + w_6)$$

$$= -w_7sech^2(w_7s_4 + w_8)sin(w_5s_3 + w_6)(w_5\frac{d}{dw_2}sin(w_3s_1 + w_4s_2))$$

$$= -w_5w_7sech^2(w_7s_4 + w_8)sin(w_5s_3 + w_6)cos(w_3s_1 + w_4s_2)\frac{d}{dw_2}(w_3s_1 + w_4s_2)$$

$$= -w_5w_7sech^2(w_7s_4 + w_8)sin(w_5s_3 + w_6)cos(w_3s_1 + w_4s_2)$$
$$(w_3\frac{d}{dw_2}tan(w_1x) + w_4\frac{d}{dw_2}sec(w_2x + s_1))$$

$$= -w_5w_7sech^2(w_7s_4 + w_8)sin(w_5s_3 + w_6)cos(w_3s_1 + w_4s_2)$$
$$(w_4sec(w_2x + s_1)tan(w_2x + s_1)\frac{d}{dw_2}(w_2x + s_1))$$

$$= -xw_4w_5w_7sech^2(w_7s_4 + w_8)sin(w_5s_3 + w_6)cos(w_3s_1 + w_4s_2)$$
$$sec(w_2x + s_1)tan(w_2x + s_1)$$

5. Differentiation of vectors/matrices

Consider vectors $\boldsymbol{u}, \boldsymbol{b} \in \mathbb{R}^d$, and matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$.

The derivative of a scalar f w.r.t. a vector $\boldsymbol{u}$ is a vector by itself, given by

$$\nabla f = \left(\frac{\partial f}{\partial u_1}, \frac{\partial f}{\partial u_2}, \ldots, \frac{\partial f}{\partial u_n}\right)$$

(**Hint**: The derivative of a scalar f w.r.t. a matrix $\boldsymbol{X}$, is a matrix whose (i,j) component is $\frac{\partial f}{\partial X_{ij}}$, where X_{ij} is the (i,j) component of the matrix $\boldsymbol{X}$.)

(a) Derive the expression for the derivative: $\nabla \boldsymbol{u}^T \boldsymbol{A} \boldsymbol{u} + \boldsymbol{b}^T \boldsymbol{u}$.

Solution: The derivation of above equation can be given as follows:

$$\nabla u^T A u + b^T u = \frac{\partial}{\partial u}(u^T A u) + \frac{\partial}{\partial u}(b^T u)$$
$$= u^T(A^T + A) + b^T$$

if matrix A is symmetric then we can write above equation as:

$$= 2u^T A + b^T$$

(b) Compare your results with derivatives for the scalar equivalents of the above expressions $au^2 + bu$.

> **Solution:** The derivation of scalar equation $au^2 + bu$ can be given as follows:
>
> $$\frac{\partial}{\partial u}(au^2 + bu) = \frac{\partial}{\partial u}(au^2) + \frac{\partial}{\partial u}(bu)$$
> $$= 2au + b$$
>
> The above obtained result is similar to the result of expression in part(a) if matrix A is symmetric.

(c) Derive the Hessian: $\frac{\partial^2 f}{\partial \mathbf{u} \partial \mathbf{u}^T}$ given that $f = \mathbf{u}^T \mathbf{A} \mathbf{u} + \mathbf{b}^T \mathbf{u}$

> **Solution:** The Hessian of the given expression can be found as follows:
>
> $$\nabla^2 u^T Au + b^T u = \frac{\partial}{\partial u}\left(\frac{\partial}{\partial u}(u^T Au + b^T u)\right)$$
> $$= \frac{\partial}{\partial u}\left(\frac{\partial}{\partial u}(u^T Au) + \frac{\partial}{\partial u}(b^T u)\right)$$
> $$= \frac{\partial}{\partial u}(u^T(A^T + A) + b^T)$$
> $$= A^T + A$$

6. **Encoding Tongue Twister :** You have been assigned a task to encode a tongue-twister phrase compactly: 'clean clams crammed in clean clans'. For convenience, you are given the frequency distribution as below.

Char	Frequency
a	5
c	5
d	1
e	3
i	1
l	4
m	3
n	4
r	1
s	2
space	5

(a) One way to encode this sequence is to use fixed length code with each code word long enough to encode ten different symbols. How many bits would be needed for this 34-character phrase using such a fixed-length code?

> **Solution:** Encode using fixed length code:
> There are ten different symbols so we'll need 4 bits to represent each character.
>
> $$\therefore total\,number\,of\,bits\,needed\,for\,34 - character\,is = 34 * 4\,bits$$
> $$= 136\,bits$$

(b) What are the minimum number of bits needed (theoretically) to encode the entire phrase , assuming that each character is independent of the surrounding character? Hint: We can calculate the average information (in other words, bits needed) of a symbol using entropy information.

> **Solution:** Encode using entropy information:
> The Shannon entropy H gives number of bits needed per symbol and is given by: $H = -\sum_k p_k \log_2 p_k$
> where, p_k is the probability of occurrence of the k-th possible symbol.
> Here, we have ten different symbols.
>
> $\therefore Average\,number\,of\,bits\,needed\,for\,a\,character\,is :$
>
> $$\begin{aligned} H = &-\sum_{k=1}^{10} p_k \log_2 p_k \\ = &-(\frac{5}{34}\log_2(\frac{5}{34}) + \frac{5}{34}\log_2(\frac{5}{34}) + \frac{1}{34}\log_2(\frac{1}{34}) + \frac{3}{34}\log_2(\frac{3}{34}) + \frac{1}{34}\log_2(\frac{1}{34}) \\ &+ \frac{4}{34}\log_2(\frac{4}{34}) + \frac{3}{34}\log_2(\frac{3}{34}) + \frac{4}{34}\log_2(\frac{4}{34}) + \frac{1}{34}\log_2(\frac{1}{34}) + \frac{2}{34}\log_2(\frac{2}{34}) \\ &+ \frac{5}{34}\log_2(\frac{5}{34})) \\ = &-(\frac{15}{34}\log_2(\frac{5}{34}) + \frac{3}{34}\log_2(\frac{1}{34}) + \frac{6}{34}\log_2(\frac{3}{34}) + \frac{8}{34}\log_2(\frac{4}{34}) + \frac{2}{34}\log_2(\frac{2}{34})) \\ = &\,3.25397 \end{aligned}$$
> $\therefore minimum\,number\,of\,bits\,needed\,to\,encode\,33 - characters\,is = \lceil 34 * (3.25397) \rceil$
> $$= \lceil 110.63498 \rceil$$
> $$= 111\,bits$$
>
> So, minimum number of bits needed to encode 33-characters is 111 bits.
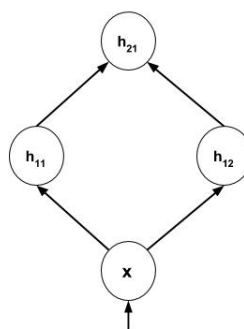
7. **Plotting Functions for Great Good**

(a) Consider the variable x and functions $h_{11}(x)$, $h_{12}(x)$ and $h_{21}(x)$ such that

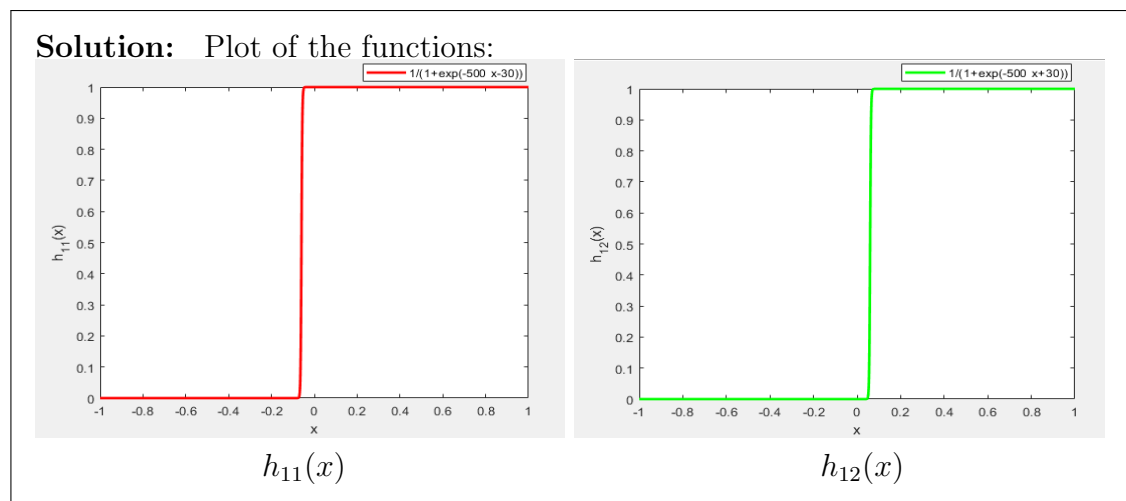$$h_{11}(x) = \frac{1}{1 + e^{-(500x+30)}}$$

$$h_{12}(x) = \frac{1}{1 + e^{-(500x-30)}}$$

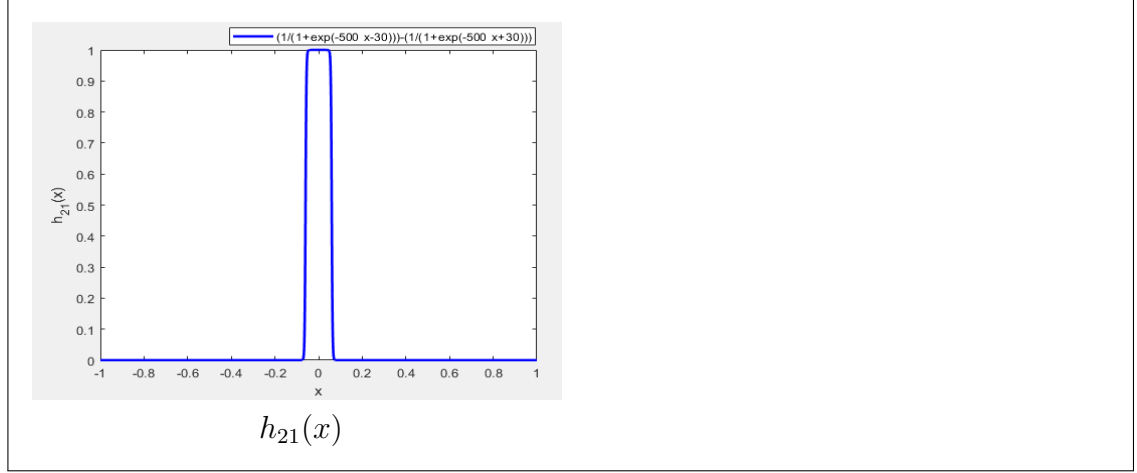$$h_{21} = h_{11}(x) - h_{12}(x)$$

The above set of functions are summarized in the graph below.

Plot the following functions: $h_{11}(x)$, $h_{12}(x)$ and $h_{21}(x)$ for $x \in (-1, 1)$

Solution: Plot of the functions:

$h_{11}(x)$

$h_{12}(x)$

$$h_{21}(x)$$

(b) Now consider the variables x_1, x_2 and the functions $h_{11}(x_1, x_2), h_{12}(x_1, x_2), h_{13}(x_1, x_2), h_{14}(x_1, x_2),$ $h_{21}(x_1, x_2), h_{22}(x_1, x_2), h_{31}(x_1, x_2)$ and $f(x_1, x_2)$ such that

$$h_{11}(x_1, x_2) = \frac{1}{1 + e^{-(x_1 + 50x_2 + 100)}}$$

$$h_{12}(x_1, x_2) = \frac{1}{1 + e^{-(x_1 + 50x_2 - 100)}}$$

$$h_{13}(x_1, x_2) = \frac{1}{1 + e^{-(50x_1 + x_2 + 100)}}$$

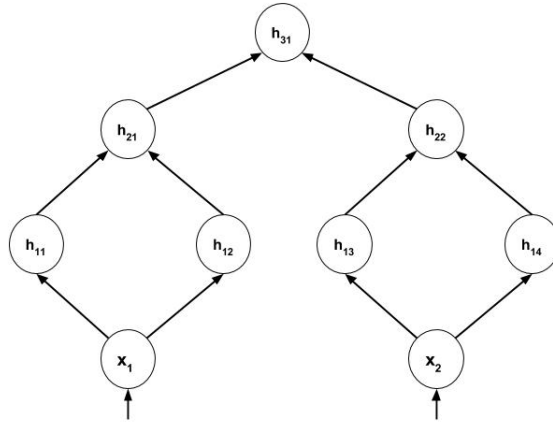$$h_{14}(x_1, x_2) = \frac{1}{1 + e^{-(50x_1 + x_2 - 100)}}$$

$$h_{21}(x_1, x_2) = h_{11}(x_1, x_2) - h_{12}(x_1, x_2)$$

$$h_{22}(x_1, x_2) = h_{13}(x_1, x_2) - h_{14}(x_1, x_2)$$

$$h_{31}(x_1, x_2) = h_{21}(x_1, x_2) + h_{22}(x_1, x_2)$$

$$f(x_1, x_2) = \frac{1}{1 + e^{-(100h_{31}(x) - 200)}}$$

The above set of functions are summarized in the graph below.

Plot the following functions: $h_{11}(x_1, x_2), h_{12}(x_1, x_2), h_{13}(x_1, x_2), h_{14}(x_1, x_2), h_{21}(x_1, x_2),$ $h_{22}(x_1, x_2), h_{31}(x_1, x_2)$ and $f(x_1, x_2)$ for $x_1 \in (-5, 5)$ and $x_2 \in (-5, 5)$

Solution: Plot of the functions: