M8-L1 Problem 1

In this problem you will solve for $\frac{\partial L}{\partial W_2}$ and $\frac{\partial L}{\partial W_1}$ for a neural network with two input features, a hidden layer with 3 nodes, and a single output. You will use the sigmoid activation function on the hidden layer. You are provided an input sample x_0, the current weights W_1 and W_2, and the ground truth value for the sample, $t = -2$

$$L = \frac{1}{2} e^T e$$

```
import numpy as np

x0 = np.array([[-2], [-6]])

W1 = np.array([[-2, 1],[3, 8],[-12, 7]])
W2 = np.array([[-11, 2, 5]])

t = np.array([[-2]])
```

Define activation function and its derivative

First define functions for the sigmoid activation functions, as well as its derivative:

```
# YOUR CODE GOES HERE
# sigmoid activation function
def sigmoid(x):
    return 1 / (1 + np.exp(-x))

# derivative of sigmoid
def sigmoid_derivative(x):
    return x * (1 - x)
```

Forward propagation

Using your activation function, compute the output of the network y using the sample x_0 and the provided weights W_1 and W_2

```
# YOUR CODE GOES HERE
# foward propagation

a1 = np.dot(W1, x0)
out1 = sigmoid(a1)
```

```python
a2 = np.dot(W2, out1)
# out2 = sigmoid(a2)
y = a2

# print the output
print("The output of the network y: ", y)
```

```
The output of the network y:  [[-1.31123207]]
```

Backpropagation

Using your calculated value of y, the provided value of t, your σ and σ' function, and the provided weights W_1 and W_2, compute the gradients $\frac{\partial L}{\partial W_2}$ and $\frac{\partial L}{\partial W_1}$.

```python
# YOUR CODE GOES HERE
# backward propagation

# derivative of the loss function with respect to the output y
f_prime = 1

# compute the gradients (with respect to weights W1 and W2)
gamma_2 = -(t - y) * f_prime
d2 = np.dot(gamma_2, out1.T)
gamma_1 = np.dot(W2.T, gamma_2) * sigmoid_derivative(out1)
d1 = np.dot(gamma_1, x0.T)

# print the gradients
print("The gradient with respect to W2:\n ", d2)
print("The gradient with respect to W1:\n ", d1)
```

```
The gradient with respect to W2:
  [[8.21031503e-02 2.43316128e-24 1.04899215e-08]]
The gradient with respect to W1:
  [[ 1.59095673e+00  4.77287018e+00]
 [-9.73264513e-24 -2.91979354e-23]
 [-1.04899214e-07 -3.14697641e-07]]
```