Sigmoid Function

Students: Pascal Baumann, Claudio Paonessa

a)

Compute the derivative of the sigmoid function $\sigma(z) = \frac{1}{1+e^{-z}}$

$$
\begin{aligned}
\frac{d}{dz}\sigma(z) &= \frac{d}{dz}\left(\frac{1}{1+e^{-z}}\right) \\
&= \frac{d}{dz}(1+e^{-z})^{-1} \\
&= -1 \cdot (1+e^{-z})^{-2} \cdot \frac{d}{dz}(e^{-z}) \\
&= -1 \cdot (1+e^{-z})^{-2} \cdot -1 \cdot (e^{-z}) \\
&= (1+e^{-z})^{-2} \cdot (e^{-z}) \\
&= \frac{e^{-z}}{(1+e^{-z})^2}
\end{aligned}
$$

b)

Show that the derivative fullfills the equation $\sigma'(z) = \sigma(z) \cdot (1 - \sigma(z))$

$$
\begin{aligned}
1 - \sigma(z) &= 1 - \frac{1}{1+e^{-z}} \\
&= \frac{1+e^{-z}}{1+e^{-z}} - \frac{1}{1+e^{-z}} \\
&= \frac{1+e^{-z}-1}{1+e^{-z}} \\
&= \frac{e^{-z}}{1+e^{-z}} \\
\sigma'(z) &= \frac{e^{-z}}{(1+e^{-z})^2} \\
&= \frac{1}{1+e^{-z}} \cdot \frac{e^{-z}}{1+e^{-z}} \\
&= \sigma(z) \cdot (1 - \sigma(z))
\end{aligned}
$$

c)

Compute the first and second derivative of $\zeta(z) = -log(\sigma(-z))$

Compute the asymptotes for $z \to \pm\infty$. Create a plot of ζ.

$$
\begin{aligned}
\frac{d}{dz} - ln(\sigma(-z)) &= \frac{d}{dz} - ln\left(\frac{1}{1+e^z}\right) \\
&= -\left(\frac{1}{1+e^z}\right)^{-1} \cdot \left(\frac{1}{(1+e^z)^2}\right) \\
&= -(1+e^z) \cdot \frac{1}{(1+e^z)^2} \\
&= -\frac{1}{1+e^z}
\end{aligned}
$$

$$
\begin{aligned}
\frac{d^2}{dz^2} - ln(\sigma(-z)) &= \frac{d}{dz} - \frac{1}{1+e^z} \\
&= \frac{d}{dz}(-1) \cdot (1+e^z)^{-1} \\
&= (-1) \cdot (-1) \cdot (1+e^z)^{-2} \cdot e^z \\
&= \frac{e^z}{(1+e^z)^2}
\end{aligned}
$$

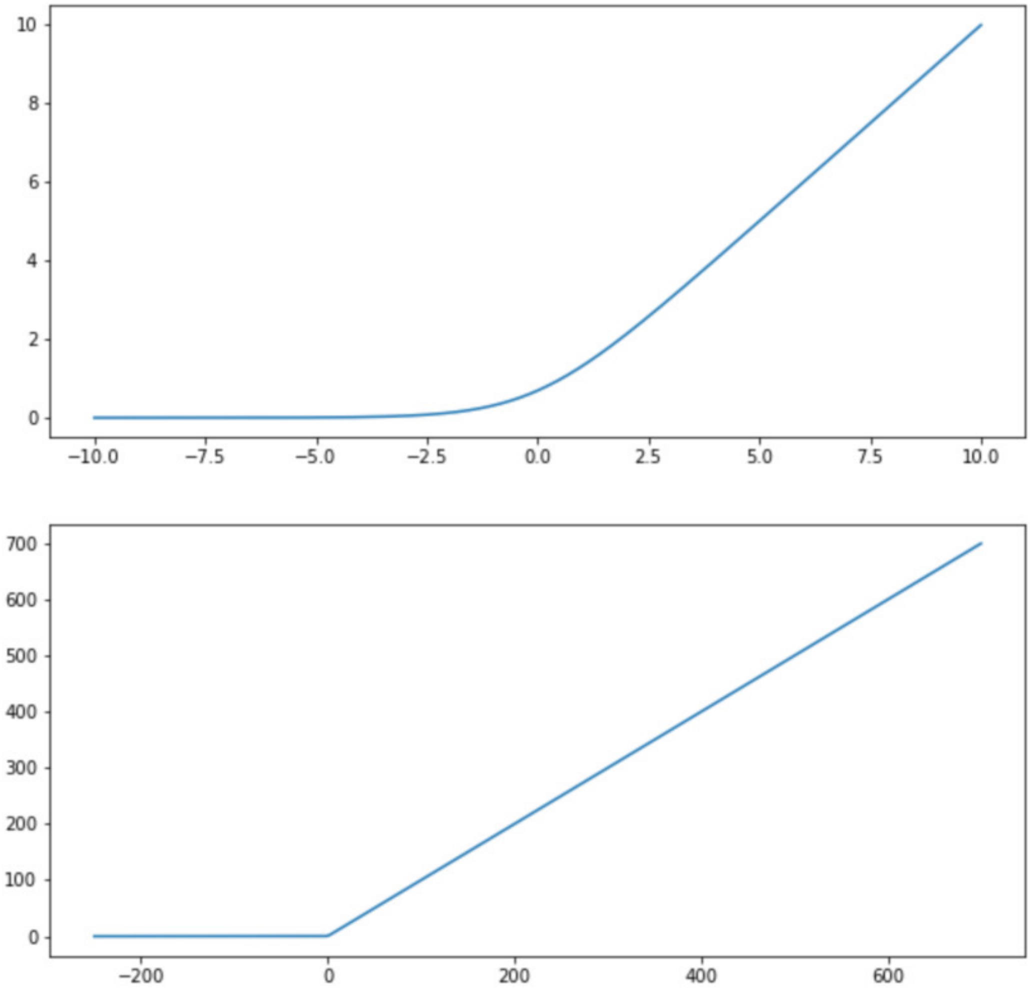
$$
\begin{aligned}
\lim_{z \to -\infty} - ln\left(\frac{1}{1+e^z}\right) &= -ln\left(\frac{1}{1+e^{-\infty}}\right) \\
&= -ln\left(\frac{1}{1+0}\right) \\
&= -ln(1) \\
&= 0 \\
\lim_{z \to -\infty} \frac{e^z}{1+e^z} &= \frac{e^{-\infty}}{1+e^{-\infty}} \\
&= 0
\end{aligned}
$$

Asymptote for minus infinity is thus the x-Axis.

$$
\begin{aligned}
\lim_{z \to \infty} - ln\left(\frac{1}{1+e^z}\right) &= -ln\left(\frac{1}{1+e^{\infty}}\right) \\
&= -ln\left(\frac{1}{1+\infty}\right) \\
&= -ln(0) \\
&= \infty \\
\lim_{z \to -\infty} \frac{e^z}{1+e^z} &= \frac{e^{\infty}}{1+e^{\infty}} \\
&= 1
\end{aligned}
$$

Asymptote for plus infinity is thus a line with an inclination of 45°

In [4]:
```
import numpy as np
import matplotlib.pyplot as plt

x = np.linspace(-10,10,10000)
y = - np.log(1/(1+np.e ** x))
plt.figure(figsize=(10,10))
plt.subplot(2,1,1)
plt.plot(x,y)
plt.subplot(2,1,2)
x = np.linspace(-250,700,10000)
y = - np.log(1/(1+np.e ** x))
plt.plot(x,y)
plt.show()
```



d)

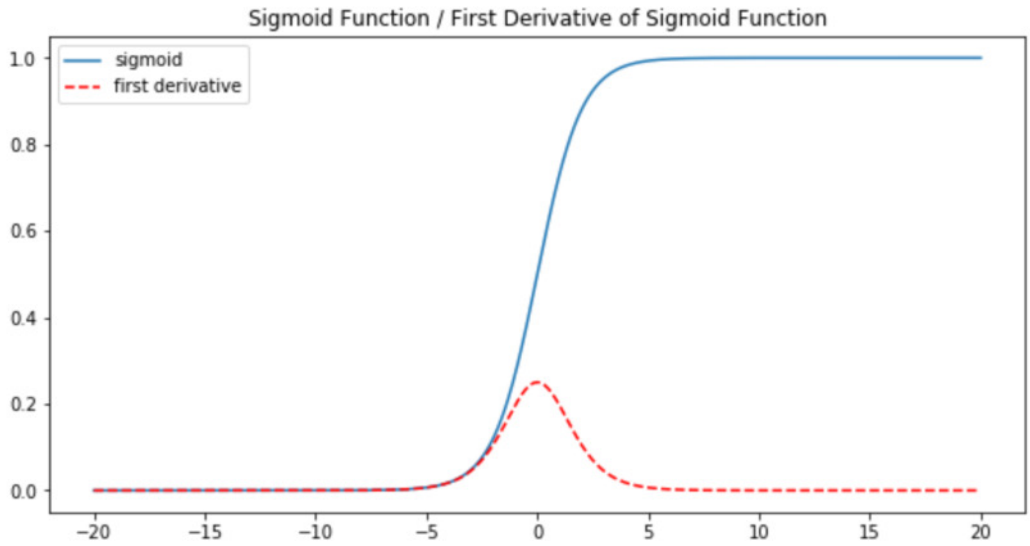
Implement the sigmoid function in a Jupyter Notebook. Make it work such that you can pass numpy arrays of arbitrary shape and the function is applied element-wise. Plot the sigmoid function and its derivative by using matplotlib.

In [22]:
```
def sigmoid(x):
    y = 1/(1 + np.e ** -x)
    return y

def first_derivative_sigmoid(x):
    y = sigmoid(x) * (1 - sigmoid(x))
    return y

def plot_sigmoid_and_derivative(x):
    y1 = sigmoid(x)
    y2 = first_derivative_sigmoid(x)
    plt.figure(figsize=(10,5))
    plt.title("Sigmoid Function / First Derivative of Sigmoid Function")
    plt.plot(x, y1, label='sigmoid')
    plt.plot(x, y2, c='r',linestyle='dashed', label='first derivative')
    plt.legend()
    plt.show()
```

In [23]:
```
x = np.array([1,2,3,4])
y = sigmoid(x)
print(y)
```

```
[0.73105858 0.88079708 0.95257413 0.98201379]
```

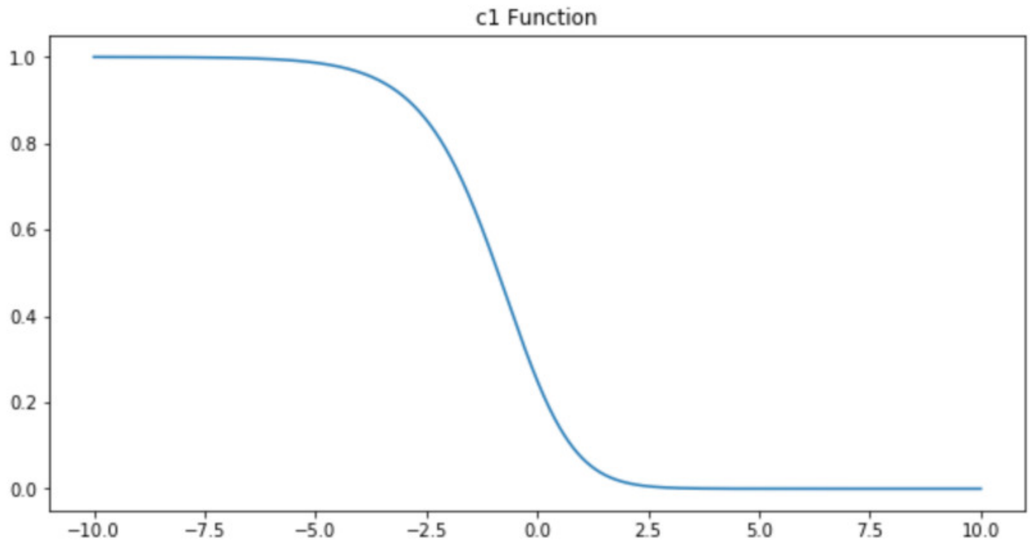
In [24]:
```
x = np.linspace(-20,20,1000)
plot_sigmoid_and_derivative(x)
```



f)

Show that the function $c_1(x) = (\sigma(x) - 1)^2$ is non-convex.

In [25]:
```python
def c1function(x):
    return (sigmoid(x)-1)**2

x = np.linspace(-10,10,1000)
plt.figure(figsize=(10,5))
plt.title("c1 Function")
plt.plot(x,c1function(x))
plt.show()
```



Add $x = -1$ the gradient becomes zero, thus we could not optimise $c_1(x)$ more by descending the gradient.

g)

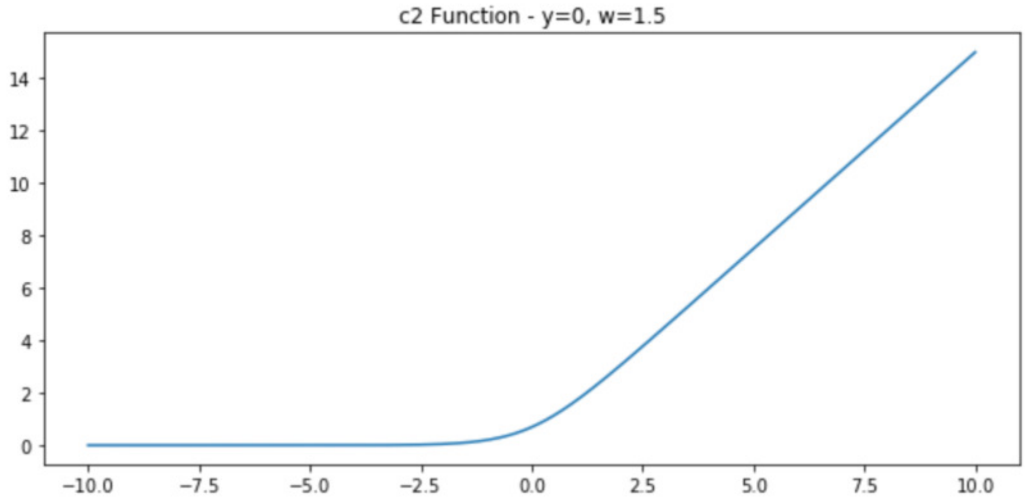
Compute the first and second derivative of the function

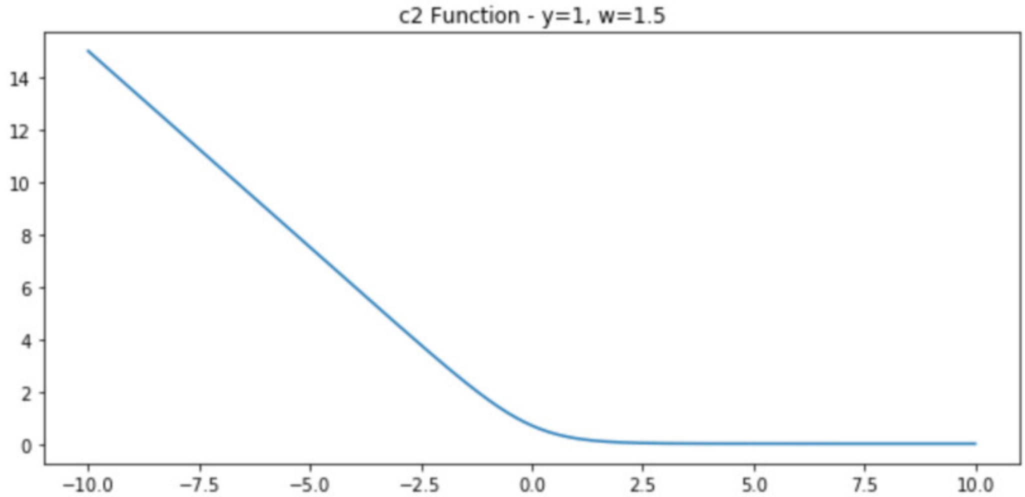
$$c_2(x) = -(y \cdot \log(\sigma(w \cdot x)) + (1-y) \cdot \log(1 - \sigma(w \cdot x)))$$

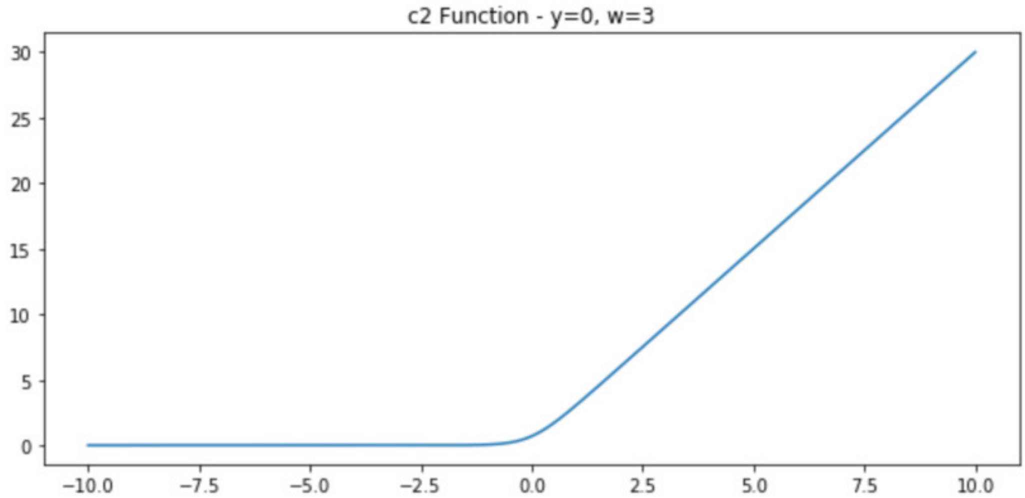
with respect to $w \in \mathbb{R}$ and for a given $y \in \{0, 1\}$. Show that c_2 is convex

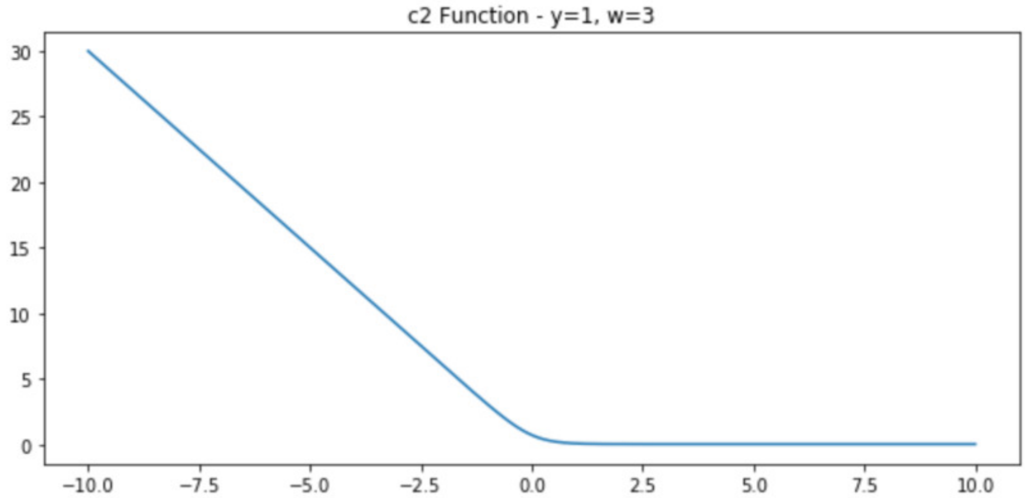
$$\frac{d}{dw}c_2 = -\frac{x((y-1)e^{wx} + y}{1 + e^{wx}}$$
$$\frac{d^2}{dw^2}c_2 = \frac{x^2 e^{wx}}{(1 + e^{wx})^2}$$

In [40]:
```
def c2function(x,y,w):
    return -(y*np.log(sigmoid(w*x)) + (1-y)* np.log(1-sigmoid(w*x)))

x = np.linspace(-10,10,1000)
plt.figure(figsize=(10,10))
plt.subplot(2,1,1)
plt.title("c2 Function - y=0, w=1.5")
plt.plot(x,c2function(x,0,1.5))
plt.subplot(2,1,2)
plt.title("c2 Function - y=1, w=1.5")
plt.plot(x,c2function(x,1,1.5))
plt.show()
plt.figure(figsize=(10,10))
plt.subplot(2,1,1)
plt.title("c2 Function - y=0, w=3")
plt.plot(x,c2function(x,0,3))
plt.subplot(2,1,2)
plt.title("c2 Function - y=1, w=3")
plt.plot(x,c2function(x,1,3))
plt.show()
```


c2 Function - y=0, w=1.5
0
2
4
6
8
10
12
14
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0


c2 Function - y=1, w=1.5
0
2
4
6
8
10
12
14
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0

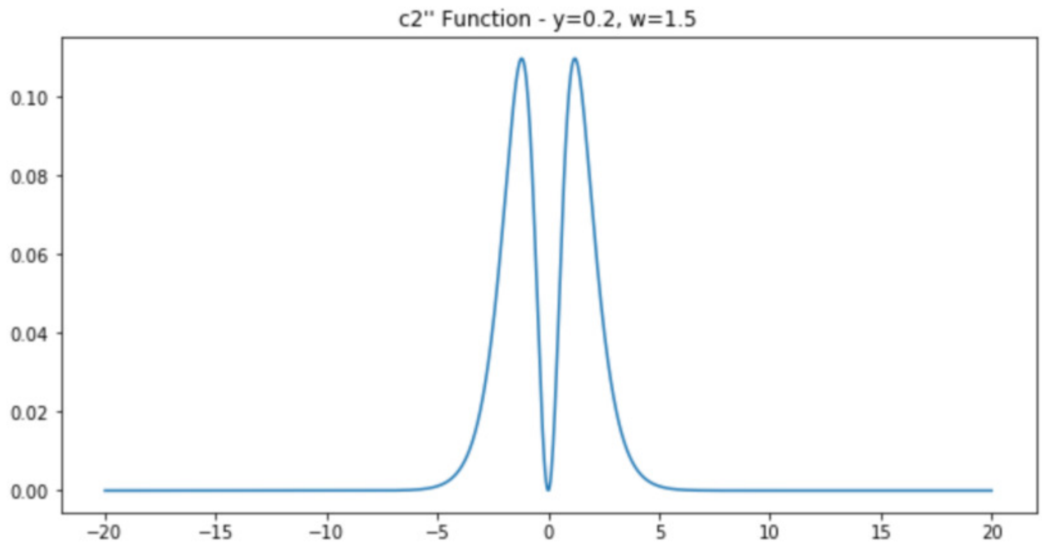




In [42]:
```
def c2function_2nd_derivative(x,y,w):
    return ((x**2) * (np.e ** (w*x)))/((np.e ** (w*x) + 1)**2)
```

In [44]:
```
x = np.linspace(-20,20,1000)
plt.figure(figsize=(10,5))
plt.title("c2'' Function - y=0.2, w=1.5")
plt.plot(x,c2function_2nd_derivative(x,0,2))
```

Out[44]: [<matplotlib.lines.Line2D at 0x16f7f6c8888>]



The second derivative with respect to w of the function c2 can not be negative. This means the function c2 is convex.

In [ ]:

```

```