Exercise

- Verify that $\sigma'(z) = \sigma(z)(1 - \sigma(z))$.

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

$$\sigma'(z) = -1\left(1+e^{-z}\right)^{-2}\left(e^{-z}\right)(-1)$$

$$= \frac{e^{-z}}{\left(1+e^{-z}\right)^{2}}$$

$$= \frac{1}{1+e^{-z}}\left(\frac{e^{-z}}{1+e^{-z}}\right)$$

$$= \frac{1}{1+e^{-z}}\left(\frac{1-1+e^{-z}}{1+e^{-z}}\right)$$

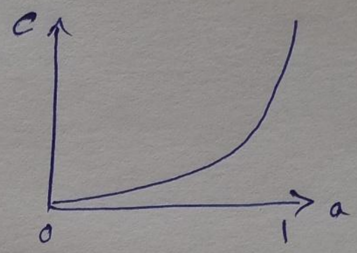
$$= \sigma(z)\left(\frac{1+e^{-z}}{1+e^{-z}} - \frac{1}{1+e^{-z}}\right)$$

$$= \sigma(z)\left(1-\sigma(z)\right)$$

- One gotcha with the cross-entropy is that it can be difficult at first to remember the respective roles of the ys and the as. It's easy to get confused about whether the right form is $-[y \ln a + (1-y)\ln(1-a)]$ or $-[a \ln y + (1-a)\ln(1-y)]$. What happens to the second of these expressions when $y = 0$ or 1? Does this problem afflict the first expression? Why or why not?

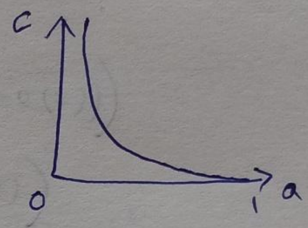
Exp 1

$C = -[y \ln a + (1-y) \ln(1-a)]$



when $y = 0$, $C = -\ln(1-a)$

$y = 1$, $C = \ln a$



Expression 2

$C = -[a \ln y + (1-a) \ln(1-y)]$

when $y = 1$, $C = \infty$ always infinity indifferent to a

when $y = 0$, $C = a \ln(0) + (1-a) * 0 = \infty$

always infinity indifferent to a

- In the single-neuron discussion at the start of this section, I argued that the cross-entropy is small if $\sigma(z) \approx y$ for all training inputs. The argument relied on y being equal to either 0 or 1. This is usually true in classification problems, but for other problems (e.g., regression problems) y can sometimes take values intermediate between 0 and 1. Show that the cross-entropy is still minimized when $\sigma(z) = y$ for all training inputs. When this is the case the cross-entropy has the value:

$$C = -\frac{1}{n} \sum_x [y \ln y + (1-y) \ln(1-y)]. \tag{64}$$

  The quantity $-[y \ln y + (1-y) \ln(1-y)]$ is sometimes known as the binary entropy.

$$C = -\left[y \ln a + (1-y) \ln(1-a)\right]$$

$$\frac{\partial C}{\partial a} = -\left[\frac{y}{a} + \frac{1-y}{1-a}(-1)\right] = \frac{1-y}{1-a} - \frac{y}{a}$$

$$= \frac{a - ay - y + ay}{a - a^2}$$

$$\frac{\partial C}{\partial a} = \frac{a-y}{a-a^2}$$

At minimum, $\frac{\partial C}{\partial a} = 0$

$$\frac{a-y}{a-a^2} = 0$$

$$\boxed{a = y}$$