Logistic/Cross-Entropy Top Gradient

$$L(y,t) = -t \ln y - (1-t)\ln(1-y)$$

$\sigma(z)$ is the logistic function.

$$\frac{\partial L}{\partial z} = \frac{\partial L}{\partial y}\frac{dy}{dz}$$ ✓ Chain rule

We know that $\frac{\partial y}{\partial z} = \sigma(z)\left(1-\sigma(z)\right) = y(1-y)$

Now for $\frac{\partial L}{\partial y}$

$$= -\frac{\partial}{\partial y}\left(t\ln y + (1-t)\ln(1-y)\right)$$

$$= -\left(\frac{t}{y} - \frac{1-t}{1-y}\right)$$

$$= -\frac{(1-y)t - y(1-t)}{y(1-y)}$$

$$= -\frac{t - yt - y + yt}{y(1-y)}$$

$$= \frac{y-t}{y(1-y)} = \frac{y-t}{\sigma'(z)}$$

✓

Put them together:

$$\frac{\partial L}{\partial z} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial z}$$

$$= \frac{y-t}{\cancel{y(1-y)}} \cancel{y(1-y)}$$

$$= y - t$$

So, then

$$\frac{\partial L}{\partial z} = y - t$$ simplest form ✓