HW3

Friday, 4 February 2022 16:06

$$\zeta = -\ln(\sigma((2y-1)z))$$

$$\frac{d\zeta}{dz} = \frac{-1}{\sigma} \cdot \sigma' \cdot (2y-1)$$

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

$$= -1 \cdot (1+e^{-z})^{-2} \cdot -1 \cdot e^{-z}$$

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

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

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

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

$$\frac{d\zeta}{dz} = \frac{-1}{\sigma} \cdot \sigma' \cdot (2y-1)$$

$$= \frac{-1 \cdot (2y-1)}{\cancel{\sigma}} \cdot \cancel{(\sigma)}(1-\sigma)$$

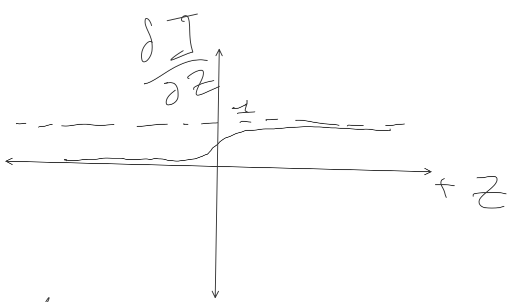
$$= (1-2y)(1-\sigma((2y-1)z))$$

$$= (1-2y)(\sigma((1-2y)z))$$

$$= \frac{1-2y}{1+\exp((2y-1)z)}$$

for $y = 0$

$$\frac{\partial J}{\partial z} = \frac{1}{1+\exp(-z)} = \sigma(z)$$

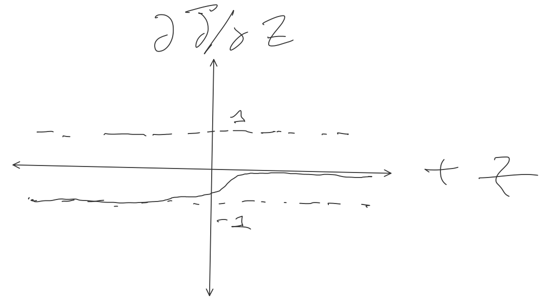


Vanishes for negative large z

For y = 0 and large negative values of z the gradient vanishes and thus the model is unable to learn

for $y = 1$

$$\frac{\partial J}{\partial z} = \frac{-1}{1+\exp(z)} = -1 \cdot \sigma(-z) = -1(1-\sigma(z)) = \sigma(z) - 1$$



Vanishes for positive large z

For y = 1 and large positive values of z the gradient vanishes and thus the model is unable to learn

This activation function has these cases where the model cannot learn and thus minimize loss as well as possible