Grafo Computacional

Son grafos dirigidos cuyos nodos se corresponden con una operación o una variable y los arcos con valores de entrada/salida

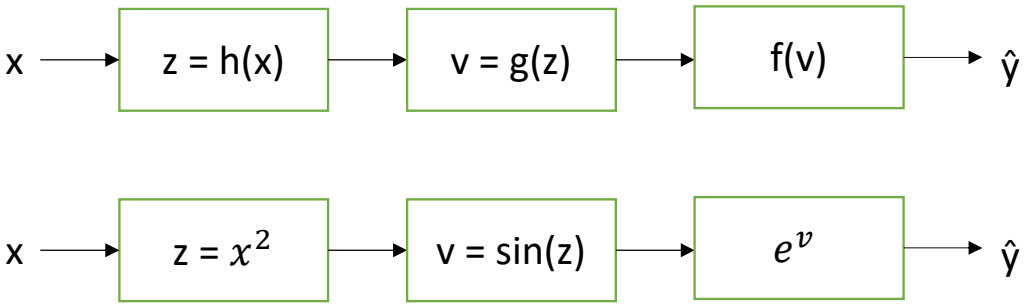
Composición de funciones

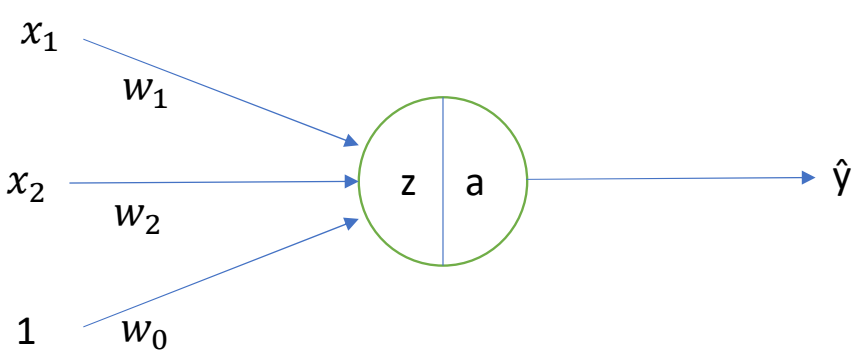
$$f\left(g\left(h(x)\right)\right) \longrightarrow e^{sin(x^2)}$$

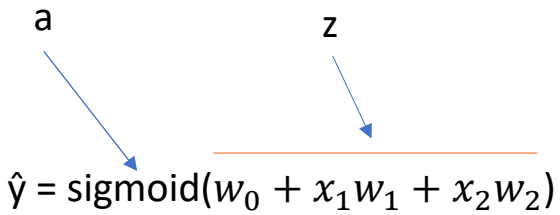
$$f(x) = e^x$$

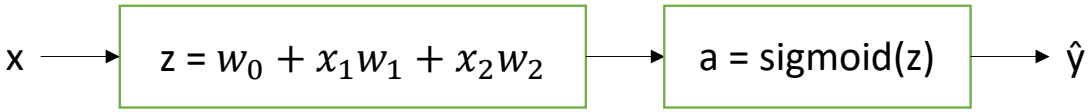
$$g(x) = sin(x)$$

$$h(x) = x^2$$




x_1
w_1
x_2
w_2
1
w_0
z
a
ŷ


a
z
ŷ = sigmoid($w_0 + x_1w_1 + x_2w_2$)


x
z = $w_0 + x_1w_1 + x_2w_2$
a = sigmoid(z)
ŷ

Descenso de gradiente

$$w_0 = w_0 - \eta \frac{dL}{dw_0}$$

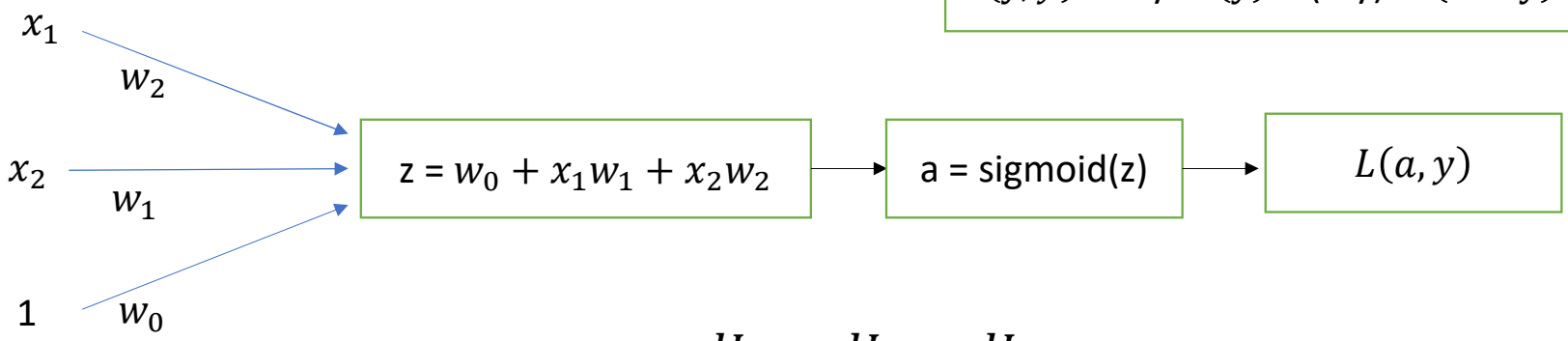
$$w_1 = w_1 - \eta \frac{dL}{dw_1}$$

$$w_2 = w_2 - \eta \frac{dL}{dw_2}$$

Regla de la Cadena

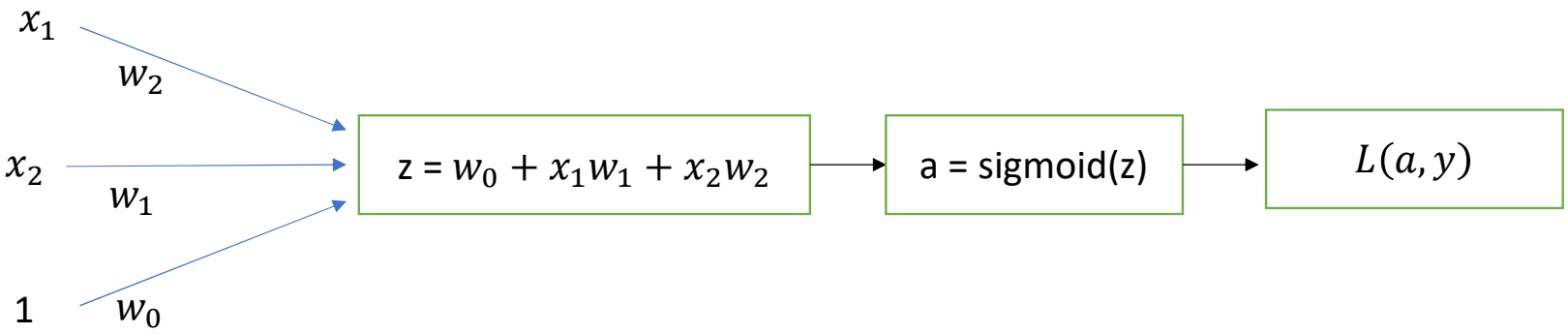
Supongamos que L sea binary cross entropy

$L(\hat{y}, y) = -\, y\, ln(\hat{y}) \; -(1-y)\, ln(1-\hat{y})$



Hay que calcular: $\frac{dL}{dw_0}$, $\frac{dL}{dw_1}$, $\frac{dL}{dw_2}$

$$\frac{dL}{dw_0} = \frac{dL}{da} \cdot \frac{da}{dz} \cdot \frac{dz}{dw_0} \qquad \frac{dL}{dw_1} = \frac{dL}{da} \cdot \frac{da}{dz} \cdot \frac{dz}{dw_1} \qquad \frac{dL}{dw_2} = \frac{dL}{da} \cdot \frac{da}{dz} \cdot \frac{dz}{dw_2}$$



$$\frac{dL}{da} = \frac{d}{da}\left(-\,y\,ln(a) - (1-y)\,ln(1-a)\right) = \frac{(1-y)}{(1-a)} - \frac{y}{a}$$

$$\frac{da}{dz} = \frac{d}{dz}\left(\frac{1}{1+e^{-z}}\right) = a\,(1-a)$$

$$\frac{dz}{dw_1} = \frac{d}{dw_1}\left(w_0 + x_1 w_1 + x_2 w_2\right) = x_1$$

$$\frac{dL}{dw_1} = \frac{dL}{da} \cdot \frac{da}{dz} \cdot \frac{dz}{dw_1} = \left(\frac{(1-y)}{(1-a)} - \frac{y}{a}\right) \cdot \left(a\,(1-a)\right) \cdot x_1$$

$$\frac{dL}{dw_1} = \frac{dL}{da} \cdot \frac{da}{dz} \cdot \frac{dz}{dw_1} = \left(\frac{(1-y)}{(1-a)} - \frac{y}{a}\right) \cdot \left(a\,(1-a)\right) \cdot x_1$$

$$= \left(\frac{(1-y)a - y(1-a)}{(1-a)a}\right) \cdot \left(a\,(1-a)\right) \cdot x_1$$

$$= \left((1-y)a - y(1-a)\right) \cdot x_1$$

$$= \left(a - ya - y + ya\right) \cdot x_1$$

$$= \left(a - y\right) \cdot x_1$$

$$w_1 = w_1 - \eta \frac{dL}{dw_1}$$

$$= w_1 - \eta\,(a-y)x_1$$

$$= w_1 - \eta\,(\hat{y}-y)x_1$$