Ejercicios de Cálculo UD5

Ejercicio 1 (Demostrar las fórmulas utilizadas para el gradiente cuando se usa la función sigmoide (1 punto).)**.**

$$\frac{\partial e}{\partial w_i} = (\hat{y} - y)\,\hat{y}\,(1 - \hat{y})\,x_i$$

$$\frac{\partial e}{\partial b} = (\hat{y} - y)\,\hat{y}\,(1 - \hat{y})$$

Solución

1.1 Definamos primero cada elemento

1.1.1 Modelo lineal

$$z = \sum_{i=1}^{n} w_i\,x_i + b$$

1.1.2 Función sigmoide

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

1.1.3 Función de error

$$e(\hat{y}) = \frac{1}{2}\,(\hat{y} - y)^2$$

1.2 Primeras derivadas

1.2.1 Derivada del modelo de datos

$$\frac{\partial z}{\partial w_i} = \lim_{h \to 0} \frac{z(w_i + h) - z(w_i)}{h}$$

Cada término para calcular el límite es:

$$z(w_i + h) = \sum_{j \neq i} w_j x_j + (w_i + h)x_i + b$$

$$z(w_i) = \sum_{j \neq i} w_j x_j + w_i x_i + b$$

Con lo que la expresión completa es:

$$\frac{z(w_i+h)-z(w_i)}{h}=\frac{\left[\sum_{j\neq i} w_j x_j+(w_i+h)x_i+b\right]-\left[\sum_{j\neq i} w_j x_j+w_i x_i+b\right]}{h}$$

Simplificando:

$$\frac{hx_i}{h}=x_i$$

Ahora el límite es:

$$\lim_{h\to 0} x_i = x_i$$

Por lo que:

$$\frac{\partial z}{\partial w_i}=x_i.$$

1.2.2 Derivada de la función sigmoide

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

Definición de derivada:

$$f'(z)=\lim_{h\to 0}\frac{f(z+h)-f(z)}{h}$$

$$f(z)=\frac{1}{1+e^{-z}},\quad f(z+h)=\frac{1}{1+e^{-(z+h)}}$$

$$f'(z)=\lim_{h\to 0}\frac{\dfrac{1}{1+e^{-(z+h)}}-\dfrac{1}{1+e^{-z}}}{h}$$

$$f'(z)=\lim_{h\to 0}\frac{(1+e^{-z})-(1+e^{-(z+h)})}{h(1+e^{-(z+h)})(1+e^{-z})}$$

$$f'(z)=\lim_{h\to 0}\frac{e^{-z}-e^{-(z+h)}}{h(1+e^{-(z+h)})(1+e^{-z})}$$

Por tanto:

$$f'(z)=\frac{e^{-z}}{(1+e^{-z})^2}=\hat{y}(1-\hat{y})$$

$$\frac{\partial \hat{y}}{\partial z}=\hat{y}\,(1-\hat{y})$$

1.2.3 Derivada de la función de error

$$e(\hat{y}) = \frac{1}{2}(\hat{y} - y)^2$$

Definición de derivada:

$$\frac{\partial e}{\partial \hat{y}} = \lim_{h \to 0} \frac{\frac{1}{2}(\hat{y} + h - y)^2 - \frac{1}{2}(\hat{y} - y)^2}{h}$$

$$= \frac{1}{2} \lim_{h \to 0} \frac{(\hat{y} - y + h)^2 - (\hat{y} - y)^2}{h}$$

Expansión del cuadrado:

$$(\hat{y} - y + h)^2 = (\hat{y} - y)^2 + 2h(\hat{y} - y) + h^2$$

$$\frac{\partial e}{\partial \hat{y}} = \frac{1}{2} \lim_{h \to 0} \frac{2h(\hat{y} - y) + h^2}{h}$$

Simplificando:

$$\frac{\partial e}{\partial \hat{y}} = \frac{1}{2} \lim_{h \to 0} [2(\hat{y} - y) + h]$$

$$\frac{\partial e}{\partial \hat{y}} = \frac{1}{2} \cdot 2(\hat{y} - y) = \hat{y} - y$$

Ahora podemos aplicar la regla de la cadena: Las derivadas que nos quedan son:

$$\frac{\partial z}{\partial w_i} = x_i$$

$$\frac{\partial \hat{y}}{\partial z} = \hat{y}(1 - \hat{y})$$

y

$$\frac{\partial e}{\partial \hat{y}} = \hat{y} - y$$

Sabemos que se trata de una composicion de funciones:

$$e = e(\hat{y}(z(w_i)))$$

Por la regla de la cadena:

$$\frac{\partial e}{\partial w_i} = \frac{\partial e}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z} \cdot \frac{\partial z}{\partial w_i}$$

Sustituyendo :

$$\frac{\partial e}{\partial \hat{y}} = \hat{y} - y$$

$$\frac{\partial \hat{y}}{\partial z} = \hat{y}(1 - \hat{y})$$

$$\frac{\partial z}{\partial w_i} = x_i$$

Entonces:

$$\boldsymbol{\frac{\partial e}{\partial w_i} = (\hat{y} - y)\, \hat{y}(1 - \hat{y})\, x_i}$$

Para b:

$$\frac{\partial e}{\partial b} = \frac{\partial e}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z} \cdot \frac{\partial z}{\partial b}$$

Pero, hay que tener en cuenta que:

$$\frac{\partial z}{\partial b} = 1$$

Por lo tanto:

$$\frac{\partial e}{\partial b} = (\hat{y} - y) \cdot \hat{y}(1 - \hat{y}) \cdot 1$$

O lo que es lo mismo::

$$\boldsymbol{\frac{\partial e}{\partial b} = (\hat{y} - y)\, \hat{y}(1 - \hat{y})}$$