

Introducción a Redes Neuronales — Tarea 1: Neuronas

David Segura, Carnet 13-11341

12/06/2020

1 Función Logística

Partiendo de la función logística

$$\varphi(v) = \frac{1}{1 + e^{-\alpha v}}$$

Aplicaremos la regla de la cadena para encontrar $\frac{\partial \varphi}{\partial v}$, quedándonos la siguiente composición de funciones:

$$f(v) = \frac{1}{v} \quad , \quad g(v) = 1 + e^{-\alpha v}$$

donde ahora,

$$\varphi(v) = f(g(v))$$

Con la aplicación de la regla de la cadena tenemos:

$$\frac{\partial \varphi}{\partial v} = \frac{\partial f(g(v))}{\partial v} \implies \frac{\partial \varphi}{\partial v} = f'(g(v))g'(v)$$

Calculando $f'(v)$ y $g'(v)$:

$$f'(v) = -\frac{1}{v^2} \quad , \quad g'(v) = -\alpha e^{-\alpha v}$$

y sustituyendo los valores de f , f' , g y g' , nos queda:

$$\frac{\partial \varphi}{\partial v} = -\frac{1}{(1 + e^{-\alpha v})^2}(-\alpha e^{-\alpha v}) = \frac{\alpha e^{-\alpha v}}{(1 + e^{-\alpha v})^2} = \alpha \left(\frac{1}{1 + e^{-\alpha v}} \right)^2 e^{-\alpha v}$$

Luego aplicando la definición de φ

$$(1) \quad \frac{\partial \varphi}{\partial v} = \alpha(\varphi(v))^2 e^{-\alpha v}$$

Por otro lado, aplicando despeje de terminos en φ ,

$$\varphi(v) = \frac{1}{1 + e^{-\alpha v}} \implies 1 + e^{-\alpha v} = \frac{1}{\varphi(v)} \implies e^{-\alpha v} = \frac{1}{\varphi(v)} - 1 \implies e^{-\alpha v} = \frac{1 - \varphi(v)}{\varphi(v)} \quad (2)$$

Finalmente, sustituyendo (2) en (1):

$$\frac{\partial \varphi}{\partial v} = \alpha(\varphi(v))^2 \left(\frac{1 - \varphi(v)}{\varphi(v)} \right) = \alpha \varphi(v) [1 - \varphi(v)]$$

2 Red Neuronal

Dada la red neuronal de dos capas en la cual la función de activación σ esta determinada por la función logística,

$$y_k = \sigma \left(\sum_{j=1}^K w_{kj}^{(2)} \sigma \left(\sum_{i=1}^M w_{ji}^{(1)} x_i \right) \right)$$

se demostrará que existe una red equivalente que calcula exactamente igual pero con función de activación en la capa oculta dada por la tangente hiperbólica.

Para ello desarrollamos la función tangente hiperbólica para encontrar la relación con la función logística:

$$\begin{aligned} \tanh \left(\alpha \frac{x}{2} \right) &= \frac{1 - e^{-\alpha v}}{1 + e^{-\alpha v}} = \frac{1}{1 + e^{-\alpha v}} - \frac{e^{-\alpha v}}{1 + e^{-\alpha v}} = \frac{1}{1 + e^{-\alpha v}} - \frac{e^{-\alpha v} + 1 - 1}{1 + e^{-\alpha v}} = \\ &= \frac{1}{1 + e^{-\alpha v}} + \frac{1}{1 + e^{-\alpha v}} - \frac{1 + e^{-\alpha v}}{1 + e^{-\alpha v}} \end{aligned}$$

y por la definición de función logística nos queda que:

$$\begin{aligned} \tanh \left(\alpha \frac{x}{2} \right) &= 2\sigma(x) - 1 \\ \implies \sigma(x) &= \frac{\tanh \left(\alpha \frac{x}{2} \right) + 1}{2} \quad (1) \end{aligned}$$

Luego, sustituyendo (1) en la capa oculta de la red neuronal inicial tenemos que

$$\begin{aligned} y_k &= \sigma \left(\sum_{j=1}^K w_{kj}^{(2)} \left(\frac{\tanh \left(\frac{\alpha}{2} \sum_{i=1}^M w_{ji}^{(1)} x_i \right) + 1}{2} \right) \right) \\ \implies y_k &= \sigma \left(\sum_{j=1}^K \frac{w_{kj}^{(2)}}{2} \tanh \left(\sum_{i=1}^M \frac{\alpha}{2} w_{ji}^{(1)} x_i \right) + \sum_{j=1}^K \frac{w_{kj}}{2} \right) \end{aligned}$$

es la red equivalente.

Podemos observar que el peso sináptico de la capa (2) ahora equivale a la mitad y el sesgo que antes era nulo en dicha capa aumentó a ser la mitad de los pesos mencionados, mientras en la capa (1) el peso cambió $\frac{\alpha}{2}$ veces.