Aprendizaje Automático: Cuestionario 2, ejercicio 5

David Cabezas Berrido

19 de mayo de 2020

1. Grafo del modelo:

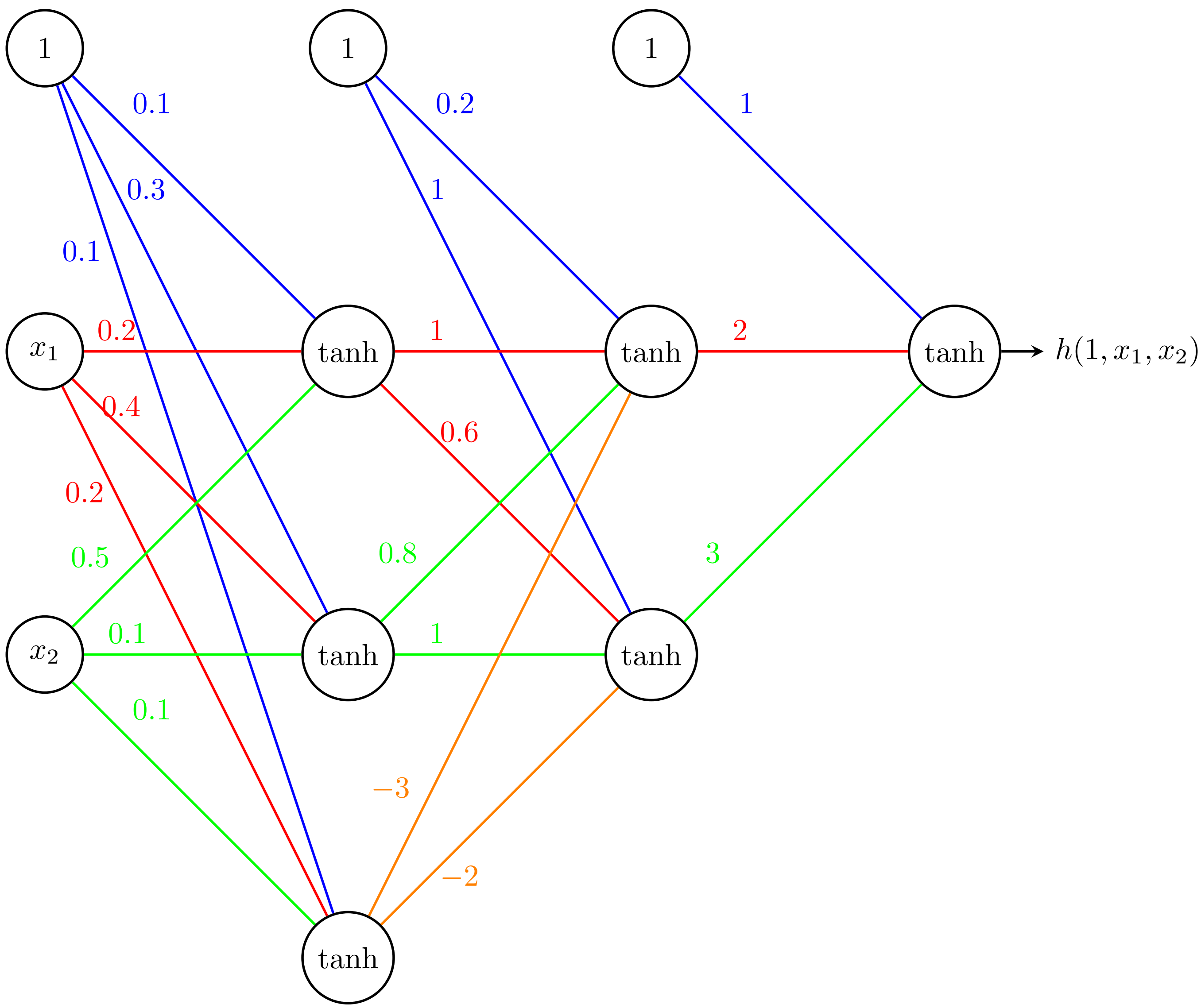

Figura 1: Grafo de la Red Neuronal

2. Propagación hacia delante

$$\mathbf{x}^{(0)} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

$$\mathbf{s}^{(1)} = W_0^T \mathbf{x}^{(0)} = \begin{pmatrix} 1{,}5 \\ 1{,}3 \\ 0{,}7 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = \begin{pmatrix} 1 \\ \tanh(\mathbf{s}^{(1)}) \end{pmatrix} = \begin{pmatrix} 1 \\ 0{,}90514825 \\ 0{,}86172316 \\ 0{,}60436778 \end{pmatrix}$$

$$\mathbf{s}^{(2)} = W_1^T \mathbf{x}^{(1)} = \begin{pmatrix} -0{,}01857655 \\ 1{,}19607656 \end{pmatrix}$$

$$\mathbf{x}^{(2)} = \begin{pmatrix} 1 \\ \tanh(\mathbf{s}^{(2)}) \end{pmatrix} = \begin{pmatrix} 1 \\ -0{,}01857441 \\ 0{,}83245396 \end{pmatrix}$$

$$\mathbf{s}^{(3)} = W_3^T \mathbf{x}^{(2)} = \begin{pmatrix} 3{,}46021305 \end{pmatrix}$$

$$\mathbf{x}^{(3)} = \begin{pmatrix} 1 \\ \tanh(\mathbf{s}^{(3)}) \end{pmatrix} = \begin{pmatrix} 1 \\ 0{,}99802713 \end{pmatrix}$$

La red precide $h(\mathbf{x}) = 0{,}99802713$ y la etiqueta es 1, así que el error es $\mathbf{e} = (h(\mathbf{x};\mathbf{w}) - y)^2 = (0{,}99802713 - 1)^2 = 3{,}8922 \cdot 10^{-6}$.

3. Propagación hacia atrás

Sensibilidades:

$$\boldsymbol{\delta}^{(3)} = \frac{\partial \mathbf{e}}{\partial \mathbf{s}^{(3)}} = 2(\mathbf{x}^{(3)} - y)\tanh'(\mathbf{s}^{(3)}) = 2(\mathbf{x}^{(3)} - y)(1 - \mathbf{x}^{(3)} \odot \mathbf{x}^{(3)}) = -1{,}555351 \cdot 10^{-5}$$

$$\boldsymbol{\delta}^{(2)} = \frac{\partial \mathbf{e}}{\partial \mathbf{s}^{(2)}} = \tanh'(\mathbf{s}^{(2)}) \odot [W^{(3)} \boldsymbol{\delta}^{(3)}]_1^2 = \begin{pmatrix} -3{,}10962877 \cdot 10^{-5} \\ -1{,}43257349 \cdot 10^{-5} \end{pmatrix}$$

$$\boldsymbol{\delta}^{(1)} = \frac{\partial \mathbf{e}}{\partial \mathbf{s}^{(1)}} = \tanh'(\mathbf{s}^{(1)}) \odot [W^{(2)} \boldsymbol{\delta}^{(2)}]_1^3 = \begin{pmatrix} -7{,}17255887 \cdot 10^{-6} \\ -1{,}00920931 \cdot 10^{-5} \\ 7{,}74003569 \cdot 10^{-5} \end{pmatrix}$$

Derivadas respecto a los errores:

$$\frac{\partial \mathbf{e}}{\partial W^{(1)}} = \mathbf{x}^{(0)} [\boldsymbol{\delta}^{(1)}]^T = \begin{pmatrix} -7{,}17255887 \cdot 10^{-6} & -1{,}00920931 \cdot 10^{-5} & 7{,}74003569 \cdot 10^{-5} \\ -1{,}43451177 \cdot 10^{-5} & -2{,}01841863 \cdot 10^{-5} & 1{,}54800714 \cdot 10^{-4} \\ -1{,}43451177 \cdot 10^{-5} & -2{,}01841863 \cdot 10^{-5} & 1{,}54800714 \cdot 10^{-4} \end{pmatrix}$$

$$\frac{\partial \mathbf{e}}{\partial W^{(2)}} = \mathbf{x}^{(1)} [\boldsymbol{\delta}^{(2)}]^T = \begin{pmatrix} -3{,}10962877 \cdot 10^{-5} & -1{,}43257349 \cdot 10^{-5} \\ -2{,}81467505 \cdot 10^{-5} & -1{,}29669139 \cdot 10^{-5} \\ -2{,}67963913 \cdot 10^{-5} & -1{,}23448176 \cdot 10^{-5} \\ -1{,}87935943 \cdot 10^{-5} & -8{,}65801257 \cdot 10^{-6} \end{pmatrix}$$

$$\frac{\partial \mathbf{e}}{\partial W^{(3)}} = \mathbf{x}^{(2)} [\boldsymbol{\delta}^{(3)}]^T = \begin{pmatrix} -1{,}55535099 \cdot 10^{-5} \\ 2{,}88897328 \cdot 10^{-7} \\ -1{,}29475809 \cdot 10^{-5} \end{pmatrix}$$