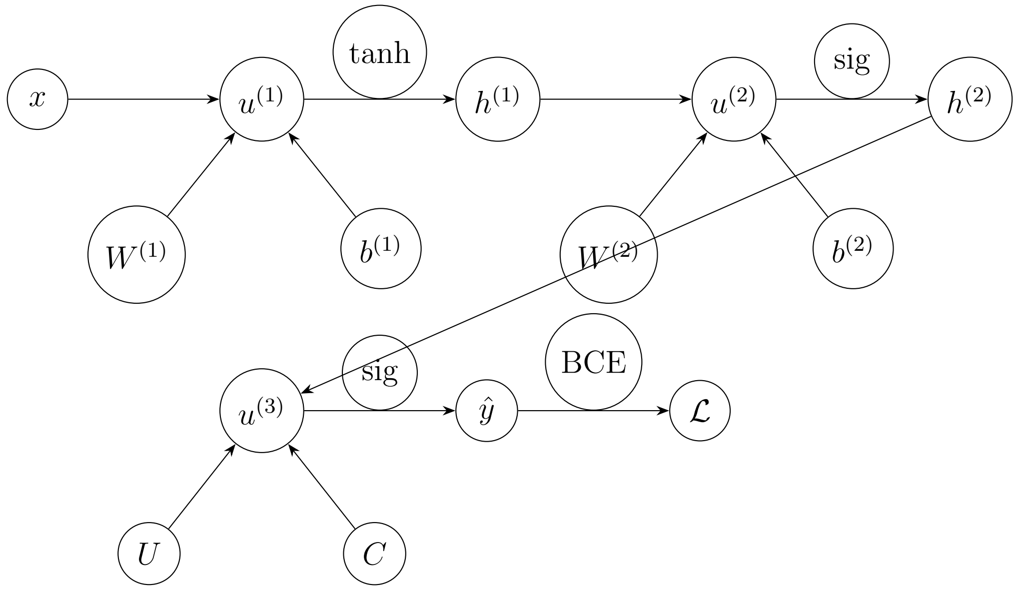
UNIVERSIDAD DE VALPARAÍSO

Instituto de Física y Astronomía
Electivo de Deep Learning

Prof. Daniela Grandón (daniela.grandon@uv.cl)

Ejemplo de Feed Forward y Backpropagation



donde la función de activación sigmoid se representará con σ.

0.1 Funciones de activación

$$\sigma(z) = \frac{1}{1+e^{-z}}, \qquad \sigma'(z) = \sigma(z)\big(1-\sigma(z)\big)$$

$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}, \qquad \tanh'(z) = 1 - \tanh^2(z)$$

0.2 Feed forward

$$u^{(1)} = xW^{(1)} + b^{(1)}, \qquad h^{(1)} = \tanh(u^{(1)})$$
$$u^{(2)} = h^{(1)}W^{(2)} + b^{(2)}, \qquad h^{(2)} = \sigma(u^{(2)})$$

$$u^{(3)} = h^{(2)}U + c, \qquad \hat{y} = \sigma(u^{(3)})$$

Función de error (entropía cruzada binaria):

$$\mathrm{BCE}\big(\hat{y}^{(i)}, y^{(i)}\big) = -\left(y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)}) \log(1 - \hat{y}^{(i)})\right)$$

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^{N} \mathrm{error}\big(\hat{y}^{(i)}, y^{(i)}\big) = H\big(P_y, P_{\hat{y}}\big)$$

0.3 Backpropagation

$$\frac{\partial \mathcal{L}}{\partial u^{(3)}} = \frac{1}{N}\big(\hat{y} - y\big)$$

$$\frac{\partial \mathcal{L}}{\partial U_i} = \sum_j \frac{\partial \mathcal{L}}{\partial u_j^{(3)}} \frac{\partial u_j^{(3)}}{\partial U_i} = \sum_j \frac{\partial \mathcal{L}}{\partial u_j^{(3)}} h_i^{(2)} \quad \Rightarrow \quad \frac{\partial \mathcal{L}}{\partial U} = \big(h^{(2)}\big)^\top \frac{\partial \mathcal{L}}{\partial u^{(3)}}$$

$$\frac{\partial \mathcal{L}}{\partial c} = \sum_j \frac{\partial \mathcal{L}}{\partial u_j^{(3)}}$$

$$\frac{\partial \mathcal{L}}{\partial h^{(2)}} = \frac{\partial \mathcal{L}}{\partial u^{(3)}} U^\top$$

$$\frac{\partial \mathcal{L}}{\partial u^{(2)}} = \frac{\partial \mathcal{L}}{\partial h^{(2)}} \odot \sigma'(u^{(2)})$$

$$\frac{\partial \mathcal{L}}{\partial W^{(2)}} = (h^{(1)})^\top \frac{\partial \mathcal{L}}{\partial u^{(2)}}, \qquad \frac{\partial \mathcal{L}}{\partial b^{(2)}} = \sum_j \frac{\partial \mathcal{L}}{\partial u_j^{(2)}}$$

$$\frac{\partial \mathcal{L}}{\partial h^{(1)}} = \frac{\partial \mathcal{L}}{\partial u^{(2)}} (W^{(2)})^\top$$

$$\frac{\partial \mathcal{L}}{\partial u^{(1)}} = \frac{\partial \mathcal{L}}{\partial h^{(1)}} \odot \tanh'(u^{(1)})$$

$$\frac{\partial \mathcal{L}}{\partial W^{(1)}} = x^\top \frac{\partial \mathcal{L}}{\partial u^{(1)}}, \qquad \frac{\partial \mathcal{L}}{\partial b^{(1)}} = \sum_j \frac{\partial \mathcal{L}}{\partial u_j^{(1)}}$$

Backpropagation con notación índicial

$$\frac{\partial \mathcal{L}}{\partial u_i^{(3)}} = \frac{1}{N}\left(\hat{y}_i - y_i\right)$$

$$\frac{\partial \mathcal{L}}{\partial U_j} = \frac{\partial \mathcal{L}}{\partial u_i^{(3)}}\frac{\partial u_i^{(3)}}{\partial U_j} = \frac{\partial \mathcal{L}}{\partial u_i^{(3)}}\, h_j^{(2)}$$

$$\frac{\partial \mathcal{L}}{\partial c} = \frac{\partial \mathcal{L}}{\partial u_i^{(3)}}$$

$$\frac{\partial \mathcal{L}}{\partial h_j^{(2)}} = \frac{\partial \mathcal{L}}{\partial u_i^{(3)}}\frac{\partial u_i^{(3)}}{\partial h_j^{(2)}} = \frac{\partial \mathcal{L}}{\partial u_i^{(3)}}\, U_j$$

$$\frac{\partial \mathcal{L}}{\partial u_{ij}^{(2)}} = \frac{\partial \mathcal{L}}{\partial h_{ij}^{(2)}}\frac{\partial h_{ij}^{(2)}}{\partial u_{ij}^{(2)}} = \frac{\partial \mathcal{L}}{\partial h_{ij}^{(2)}}\, \sigma'\Big(u_{ij}^{(2)}\Big)$$

$$\frac{\partial \mathcal{L}}{\partial W_{kl}^{(2)}} = \frac{\partial \mathcal{L}}{\partial u_{il}^{(2)}}\frac{\partial u_{il}^{(2)}}{\partial W_{kl}^{(2)}} = \frac{\partial \mathcal{L}}{\partial u_{il}^{(2)}}\, h_k^{(1)}$$

$$\frac{\partial \mathcal{L}}{\partial b_l^{(2)}} = \frac{\partial \mathcal{L}}{\partial u_{il}^{(2)}}$$

$$\frac{\partial \mathcal{L}}{\partial h_k^{(1)}} = \frac{\partial \mathcal{L}}{\partial u_{il}^{(2)}}\frac{\partial u_{il}^{(2)}}{\partial h_k^{(1)}} = \frac{\partial \mathcal{L}}{\partial u_{il}^{(2)}}\, W_{kl}^{(2)}$$

$$\frac{\partial \mathcal{L}}{\partial u_{ik}^{(1)}} = \frac{\partial \mathcal{L}}{\partial h_k^{(1)}}\frac{\partial h_k^{(1)}}{\partial u_{ik}^{(1)}} = \frac{\partial \mathcal{L}}{\partial h_k^{(1)}}\, \tanh'\Big(u_{ik}^{(1)}\Big)$$

$$\frac{\partial \mathcal{L}}{\partial W_{fk}^{(1)}} = \frac{\partial \mathcal{L}}{\partial u_{ik}^{(1)}}\frac{\partial u_{ik}^{(1)}}{\partial W_{fk}^{(1)}} = \frac{\partial \mathcal{L}}{\partial u_{ik}^{(1)}}\, x_f$$

$$\frac{\partial \mathcal{L}}{\partial b_k^{(1)}} = \frac{\partial \mathcal{L}}{\partial u_{ik}^{(1)}}$$