

$$= \frac{e^{-\theta^{(2)} a^{(2)}}}{(1 + e^{-\theta^{(2)} a^{(2)}})^2} = \phi$$

Logo:  $g'(z^{(3)}) = a^{(3)} \cdot (1 - a^{(3)})$

mudando de escala:

Então, voltando à análise dimensional:

~~$\Theta^{(3)}$  has size  $6 \times 4$~~  → Nemos tb considerar  
 ~~$a^{(3)}$  has size  $6 \times 5000$~~   $\delta$  para a  
 ~~$(1 - a^{(3)})$  has size  $6 \times 5000$~~  lias unit  
 mentira! Na BP,  
 não Nemos, mas.

$$\text{Então: } (a^{(3)})^T \cdot (1 - a^{(3)}) = N_{5000 \times 5000}$$

→ it should be element-wise multiplication!

$$\text{Então: } a^{(3)} \cdot (1 - a^{(3)}) = N_{6 \times 5000}$$

OBS: começo a perceber que vai ser importante armazenar todos os  $a^{(l)}$  em uma matriz  $A^{(l)}$  ( $\max(s(l)) \times L$ )' l, como  $s(l)$  desconsidera lias unit, está correto.

Portanto:

$$\delta^{(3)}_4 = (\theta^{(3)})^T \cdot g^{(4)} \cdot a^{(3)} \cdot (1 - a^{(3)})$$

para  $\theta^{(3)}$  →  $6 \times 4$   $4 \times 5000$   $6 \times 5000$   $6 \times 5000$   
 também  $6 \times 5000$   
 não considerarmos  $6 \times 5000$   
 lias unit na BP.

$$\delta^{(3)}_{6 \times 5000}$$

(isso pq a 1ª coluna de  $\theta^{(3)}$ , que se refere aos lias unit, não é desconsiderada)

Além disso:  $\delta^{(2)} = (\theta^{(2)})^T \cdot S^{(3)} \cdot a^{(2)} \cdot (1 - a^{(2)})$

oops!

mismatch!

UO

OBS2:  $\delta = a - s$ : estamos dizendo que  $\delta$  é o "erro" do valor de ativação, não de  $\Theta$ . Dessa forma, como sempre temos  $a_0^{(l)} = 1$  (lias unit), me parece ser razoável desconsiderar  $\delta_0^{(l)}$  para os termos de lias unit.

OBS3: mesmo assim, precisaríamos atualizar todos os valores  $\theta_{j,0}^{(l)}$ .

Isso é o que  $\delta$  faz!