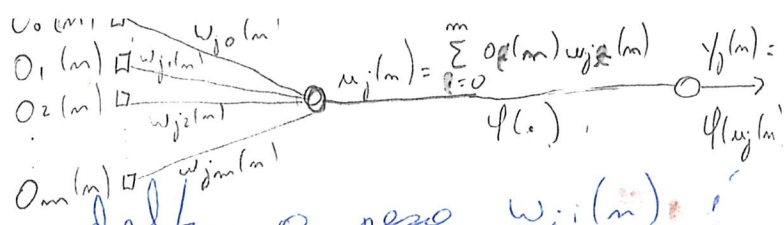


# Back propagation



De acordo com a regra delta, o peso  $w_{ji}(m)$  ajustado por

$$\Delta w_{ji}(m) = -\eta \frac{\partial E(m)}{\partial w_{ji}(m)} \quad (1)$$

De acordo com a regra da cadeia do cálculo, podemos reescrever (1) como

$$\Delta w_{ji}(m) = -\eta \frac{\partial E(m)}{\partial x_j(m)} \frac{\partial x_j(m)}{\partial u_j(m)} \frac{\partial u_j(m)}{\partial w_{ji}(m)} \quad (2)$$

Diferenciando a Eq. ~~(2.30)~~ <sup>(4.1)</sup> que define a saída do neurônio  $j$  em relação a  $u_j(m)$ , temos

$$\frac{\partial x_j(m)}{\partial u_j(m)} = \frac{\partial (\phi(u_j(m)))}{\partial u_j(m)} = \phi'(u_j(m)) \quad (3)$$

Diferenciando a Eq. ~~(2.31)~~ <sup>(4.2)</sup> que define a ativação interna do neurônio  $j$  em relação a  $w_{ji}(m)$ , temos

$$\frac{\partial u_j(m)}{\partial w_{ji}(m)} = \frac{\partial \left( \sum_{l=0}^m w_{jl}(m) o_l(m) \right)}{\partial w_{ji}(m)} = o_i(m) \quad (4)$$

Substituindo (4) e (3) em (2), então

$$\Delta w_{ji}(n) = -\eta \frac{\partial E(n)}{\partial y_{ji}(n)} \cdot \psi'(u_{ji}(n)) \cdot o_i(n)$$

$$\text{ou } \boxed{\Delta w_{ji}(n) = \eta \cdot \delta_j(n) \cdot o_i(n)} \quad (5)$$

se definirmos

$$\delta_j(n) = -\frac{\partial E(n)}{\partial y_{ji}(n)} \psi'(u_{ji}(n)) \quad (6)$$

Agora as coisas podem ser consideradas.

Caso 1: O neurônio j está na camada de saída.

Neste caso, podemos usar novamente a regra do cadeia para escrever:

$$\frac{\partial E(n)}{\partial y_{ji}(n)} = \frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_{ji}(n)} \quad (7)$$

Diferenciando ambos os lados da Eq. <sup>(4.6)</sup>~~(2.35)~~, que define a energia média do erro quadrático em relação a  $e_j(n)$ , obtemos

$$\frac{\partial E(n)}{\partial e_j(n)} = \frac{\partial \left( \frac{1}{2} \sum_{k \in c} e_k^2(n) \right)}{\partial e_j(n)} = e_j(n) \quad (8)$$

Diferenciando a Eq. <sup>(4.5)</sup> ~~(2.34)~~, que define o erro instantâneo na saída  $j$ , em relação a  $y_j(n)$ , temos

$$\frac{\partial e_j(n)}{\partial y_j(n)} = \frac{\partial (d_j(n) - y_j(n))}{\partial y_j(n)} = -1 \quad (9)$$

Substituindo (9) e (8) em (7), obtemos

$$\frac{\partial E(n)}{\partial y_j(n)} = -e_j(n) \quad (10)$$

Substituindo (10) em (6) obtemos

$$\boxed{\delta_j(n) = e_j(n) \psi'(u_j(n))} \quad (11)$$

que define o gradiente local quando o neurônio  $j$  está na camada de saída.

Caso 2: O neurônio  $j$  está em uma camada oculta

Repare que neste caso, não existe um resposta desejada para compararmos com a saída do neurônio. Neste caso, o sinal de erro para um neurônio em uma camada oculta deve ser determinado recursivamente.

Seja  $n$  ~~um~~ o índice dos neurônios da camada de saída.

Neste caso, podemos reescrever a Eq. ~~(2.35)~~ <sup>4.6</sup> como  
(12)

$$E(m) = \frac{1}{2} \sum_{n \in C} e_n^2(m)$$

Diferenciando (12) em relação a saída do neurônio  $j$ , obtemos

$$\frac{\partial E(m)}{\partial y_j(m)} = \sum_n e_n \frac{\partial e_n(m)}{\partial y_j(m)} \quad (13)$$

Usando a regra da cadeia, podemos reescrever (13) como

$$\frac{\partial E(m)}{\partial y_j(m)} = \sum_n e_n(m) \frac{\partial e_n(m)}{\partial u_n(m)} \frac{\partial u_n(m)}{\partial y_j(m)} \quad (14)$$

Como o erro em cada  $n$  é de saída e é dado por ~~(Eq. 2.34)~~ <sup>(4.5)</sup>

$$e_n(m) = d_n(m) - y_n(m)$$

ou

$$e_n(m) = d_n(m) - \phi(u_n(m)) \quad (15)$$

~~então (14) torna-se~~

$$\frac{\partial E(m)}{\partial y_j(m)} = \sum_n e_n(m) (-\phi'(u_n(m)))$$

então:

$$\frac{\partial e_n(m)}{\partial u_n(m)} = -\phi'(u_n(m)) \quad (16)$$

4.2

Além disso, usando a eq. ~~(2.51)~~ para calcular a soma do numerador  $k$  da camada de saída,

então

$$\frac{\partial u_k(m)}{\partial y_j(m)} = \frac{\partial \left( \sum_{l=0}^m w_{kl}(m) \phi_l(m) \right)}{\partial y_j(m)} = \frac{\partial \left( \sum_{l=0}^m w_{kl}(m) y_l(m) \right)}{\partial y_j(m)}$$

$$\frac{\partial u_k(m)}{\partial y_j(m)} = w_{kj}(m) \quad (17)$$

Substituindo (16) e (17) em (14), obtemos

$$\frac{\partial E(m)}{\partial y_j(m)} = - \sum_n e_n(m) \cdot \phi'(u_n(m)) \cdot w_{nj}(m)$$

$$\frac{\partial E(m)}{\partial y_j(m)} = - \sum_n \delta_n(m) w_{nj}(m) \quad (18)$$

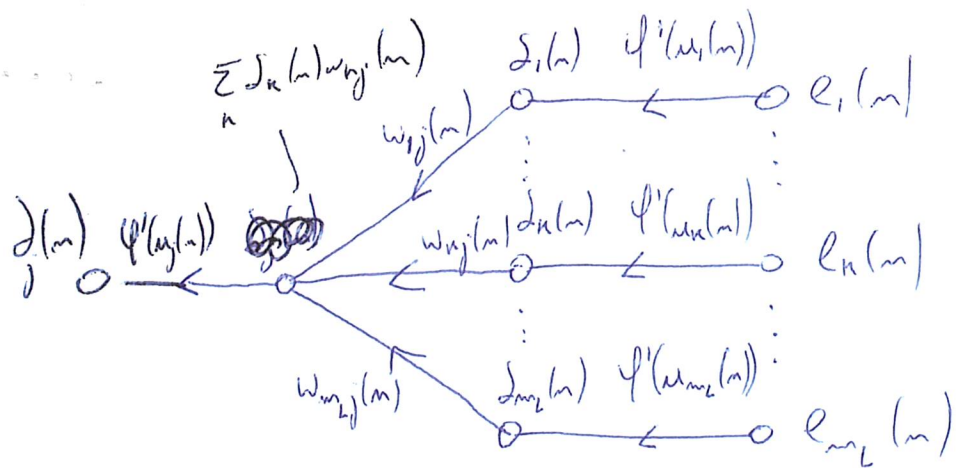
seja o gradiente local  $\delta_n(m)$  dado por (11)

Finalmente substituindo (18) em (6), temos

que para o neurônio  $j$  da camada oculta:

$$\boxed{\delta_j(m) = \phi'(u_j(m)) \sum_n \delta_n(m) w_{nj}(m)} \quad (19)$$

Adicionar a seguir mostra a representação por grafo de fluxo de sinal de (19)



## Derivada da Função Logística

Sabe-se que a ~~função~~ <sup>função</sup> logística é dada por

$$\psi(u_j(n)) = \frac{1}{1 + e^{-au_j(n)}} \quad (1)$$

para  $a > 0$  e  $-\infty < u_j(n) < +\infty$

Resposta: ~~Diferencial~~ <sup>Diferencial</sup> (1) em relação a  $u_j(n)$ .

Fazendo uma mudança de variável:

$$z = 1 + e^{-au_j(n)} \quad (2)$$

tais que

$$\psi^z(z) = \frac{1}{z} = z^{-1} \quad (3)$$

que derivando em relação a  $z$ , resulta em

$$\psi'(z) = -1 z^{-2} dz \quad (4)$$

substituindo (2) em (4)

$$\psi'(u_j(n)) = -1 \left( \frac{1}{1 + e^{-au_j(n)}} \right)^2 (-a e^{-au_j(n)})$$

$$\psi'(u_j(n)) = \frac{a e^{-au_j(n)}}{(1 + e^{-au_j(n)})^2} \quad (5)$$

Podemos manipular (5) de tal forma que

$$\psi'(u_j(n)) = a \frac{1}{1 + e^{-au_j(n)}} \frac{e^{-au_j(n)}}{1 + e^{-au_j(n)}} = a \frac{1}{1 + e^{-au_j(n)}} \left( 1 - \frac{1}{1 + e^{-au_j(n)}} \right)$$

Considerando que:

$$y_j(n) = \psi(w_j(n)) = \frac{1}{1 + e^{-aw_j(n)}}$$

(7)

Então, substituindo (7) em (6)

$$\psi'(w_j(n)) = a y_j(n) (1 - y_j(n))$$

(8)