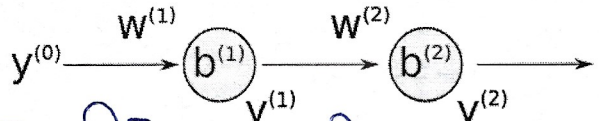
Bartłomiej Kowalewski, 195204
Wtorek 15:10

EIO – Laboratorium 2

# Algorytm wstecznej propagacji błędu

## Zadanie 1

**Schemat sieci neuronowej:**




$\frac{\partial E}{\partial w_2} = \frac{\partial E}{\partial y_2} \cdot \frac{\partial y_2}{\partial z_2} \cdot \frac{\partial z_2}{\partial w_2} = 2(y_2 - \hat{y}) \cdot RELU'(z_2) \cdot y_1$

**Zbiór uczący:**

$X_1 = (0, \hat{y} = 0)$
$X_2 = (4, \hat{y} = 2)$

**Funkcja aktywacji:** ReLU

**Funkcja błędu:** MSE

a) Korzystając z reguły łańcuchowej wyprowadź wzory na obliczenie pochodnych cząstkowych:

$\frac{\partial E}{\partial w^{(2)}} = \frac{\partial E}{\partial y_2} \cdot \frac{\partial y_2}{\partial z_2} \cdot \frac{\partial z_2}{\partial w_2} = 2(y_2 - \hat{y}) \cdot RELU'(z_2) \cdot y_1$

$\frac{\partial E}{\partial w^{(1)}} = \frac{\partial E}{\partial y_2} \cdot \frac{\partial y_2}{\partial z_2} \cdot \frac{\partial z_2}{\partial y_1} \cdot \frac{\partial y_1}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_1} = 2(y_2 - \hat{y}) \cdot RELU'(z_2) \cdot w_2 \cdot RELU'(z_1) \cdot y_0$

$\frac{\partial E}{\partial b^{(2)}} = \frac{\partial E}{\partial y_2} \cdot \frac{\partial y_2}{\partial z_2} \cdot \frac{\partial z_2}{\partial b_2} = 2(y_2 - \hat{y}) \cdot RELU'(z_2) \cdot 1$

$\frac{\partial E}{\partial b^{(1)}} = \frac{\partial E}{\partial y_2} \cdot \frac{\partial y_2}{\partial z_2} \cdot \frac{\partial z_2}{\partial y_1} \cdot \frac{\partial y_1}{\partial z_1} \cdot \frac{\partial z_1}{\partial b_1} = 2(y_2 - \hat{y}) \cdot RELU'(z_2) \cdot w_2 \cdot RELU'(z_1) \cdot 1$

b) Oblicz zaktualizowane wagi dla $\mu = 0.1$ oraz następujących wartości wag: $w^{(1)} = 0.2, b^{(1)} = 0.5, w^{(2)} = 0.5, b^{(2)} = -0.5$

$w_1' = w_1 - \mu \frac{\partial E}{\partial w_1} = w_1 - \mu \frac{1}{n} \sum_{i=1}^{n} \frac{\partial E(x_i)}{\partial w_i}$

$d_1 = 2(0 - 0 \cdot 0) \cdot 0{,}5 \cdot 1 \cdot 0 = 0$

$X_1 = (0, 0)$
$z_1 = 0 \cdot 0{,}2 + 0{,}5 = 0{,}5$
$y_1 = RELU(z_1) = RELU(0{,}5) = 0{,}5$
$z_2 = 0{,}5 \cdot 0{,}5 - 0{,}5 = -0{,}25$
$y_2 = RELU(-0{,}25) = 0$
$MSE = (0 - 0)^2 = 0$

$d_2 = 2(0{,}15 - 2) \cdot 1 \cdot 0{,}5 \cdot 4 =$

$d_1 = 0$
$d_2 = -7{,}4$ $= 2 \cdot (-1{,}85) \cdot 1 \cdot 2 =$

$4 \cdot (-1{,}85) =$

$= (-7{,}4)$

dla $w_1'$

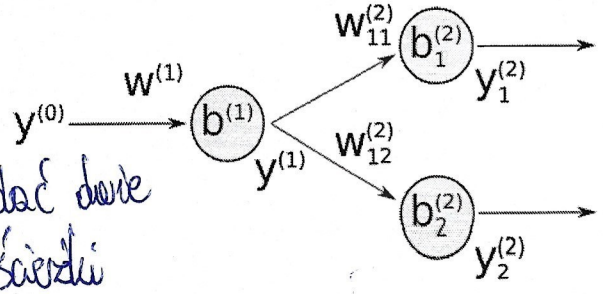
$d_1 = \frac{\partial E}{\partial w_1}(x_1), d_2 = \frac{\partial E}{\partial w_1}(x_2)$

$X_2 = (4, 2)$
$z_1 = 0{,}2 \cdot 4 + 0{,}5 = 1{,}3$
$y_1 = RELU(1{,}3) = 1{,}3$
$z_2 = 0{,}5 \cdot 1{,}3 - 0{,}5 = 0{,}15$
$y_2 = RELU(0{,}15) = 0{,}15$
$w_1' = w_1 - \mu \cdot \frac{1}{2} \cdot (d_1 + d_2)$

$w_1' = w_1 - \mu \cdot \frac{1}{2} \cdot (d_1 + d_2) =$
$= 0{,}2 - 0{,}1 \cdot 0{,}5 \cdot (0 + (-7{,}4)) =$
$= 0{,}57$

## Zadanie 2

**Schemat sieci neuronowej:**




dodać dwie ścieżki

**Zbiór uczący:**

$X_1 = (0, \hat{y} = (0, 1))$
$X_2 = (4, \hat{y} = (2, 5))$

**Funkcja aktywacji:** ReLU

**Funkcja błędu:** MSE

a) Korzystając z reguły łańcuchowej wyprowadź wzory na obliczenie pochodnych cząstkowych:

$$\frac{\partial E}{\partial w_{11}^{(2)}} = \frac{\partial E}{\partial y_1^{(2)}} \cdot \frac{\partial y_1^{(2)}}{\partial z_1^{(2)}} \cdot \frac{\partial z_1^{(2)}}{\partial w_{11}^{(2)}} = 2(y_1^{(2)} - \hat{y}) \cdot RELU'(z_1^{(2)}) \cdot y^{(1)}$$

$$\frac{\partial E}{\partial w_{12}^{(2)}} = \frac{\partial E}{\partial y_2^{(2)}} \cdot \frac{\partial y_2^{(2)}}{\partial z_2^{(2)}} \cdot \frac{\partial z_2^{(2)}}{\partial w_{12}^{(2)}} = 2(y_2^{(2)} - \hat{y}) \cdot RELU'(z_2^{(2)}) \cdot y^{(1)}$$

$$\frac{\partial E}{\partial w^{(1)}} = \frac{\partial E}{\partial y_1^{(2)}} \cdot \frac{\partial y_1^{(2)}}{\partial z_1^{(2)}} \cdot \frac{\partial z_1^{(2)}}{\partial y^{(1)}} \cdot \frac{\partial y^{(1)}}{\partial z^{(1)}} \cdot \frac{\partial z^{(1)}}{\partial w^{(1)}} + \frac{\partial E}{\partial y_2^{(2)}} \cdot \frac{\partial y_2^{(2)}}{\partial z_2^{2}} \cdot \frac{\partial z_2^{(2)}}{\partial y^{(1)}} \cdot \frac{\partial y^{(1)}}{\partial z^{(1)}} \cdot \frac{\partial z^{(1)}}{\partial w^{(1)}} =$$

$$\frac{\partial E}{\partial b^{(1)}} = \left[2(y_1^{(2)} - y) \cdot RELU'(z_1^{(2)}) \cdot w_1^{(2)} + 2(y_2^{(2)} - y) \cdot RELU'(z_2^{(2)}) \cdot w_2^{(2)}\right] \cdot RELU'(z^{(1)}) \cdot y^0$$

Podobnie jak w poprzednim przypadku $\left(\frac{\partial E}{\partial w^{(1)}}\right)$, wynik → $\left[2(y_1^{(2)} - y) \cdot RELU'(z_1^{(2)}) \cdot w_1^{(2)} + 2(y_2^{(2)} - y) \cdot RELU'(z_2^{(2)}) \cdot w_2^{(2)}\right] \cdot RELU'(z^{(1)}) \cdot 1$

b) Oblicz zaktualizowane wagi dla $\mu = 0.1$ oraz następujących wartości wag: $w^{(1)} = 0.2, b^{(1)} = 0.5, w_{11}^{(2)} = 0.5, b_1^{(2)} = -0.5, w_{12}^{(2)} = 0.0, b_2^{(2)} = 0.5$

$X_1 = (0, (0,1))$

$z^{(1)} = 0 \cdot 0{,}2 + 0{,}5 = 0{,}5$

$y^{(1)} = RELU(z_1) = 0{,}5$

$z_1^{(2)} = 0{,}5 \cdot 0{,}5 - 0{,}5 = -0{,}25$

$y_1^{(2)} = RELU(-0{,}25) = 0$

$z_2^{(2)} = 0 \cdot 0{,}5 + 0{,}5 = 0{,}5$

$y_2^{(2)} = RELU(0{,}5) = 0{,}5$

$w_{11}^{(2)'} = w_{11}^{(2)} - \mu \cdot \frac{1}{2}(d_1 + d_2)$

$d_1 = \frac{\partial E}{\partial w_{11}^{(2)}}(X_1) = 2 \cdot (0 - 0) \cdot 0 \cdot 0{,}5 = \boxed{0}$

$d_2 = \frac{\partial E}{\partial w_{11}^{(2)}}(X_2) = 2 \cdot (0{,}15 - 2) \cdot 1 \cdot 1{,}3 = (-4{,}81)$

$w_{11}^{(2)'} = 0{,}5 - 0{,}1 \cdot 0{,}5 \cdot (0 + (-4{,}81)) = \boxed{0{,}7405}$

$X_2 = (4, (2,5))$

$z^{(1)} = 0{,}2 \cdot 4 + 0{,}5 = 1{,}3$

$y^{(1)} = RELU(1{,}3) = 1{,}3$

$z_1^{(2)} = 0{,}5 \cdot 1{,}3 - 0{,}5 = 0{,}15$

$y_1^{(2)} = RELU(0{,}15) = 0{,}15$

$z_2^{(2)} = 0 \cdot 1{,}3 + 0{,}5 = 0{,}5$

$y_2^{(2)} = RELU(0{,}5) = 0{,}5$

$w_{12}^{(2)'} = w_{12}^{(2)} - \mu \cdot \frac{1}{2}(d_1 + d_2)$

$d_1 = \frac{\partial E}{\partial w_{12}^{(2)}}(X_1) = 2 \cdot (0{,}5 - 1) \cdot 1 \cdot 0{,}5 = (-0{,}5)$

$d_2 = \frac{\partial E}{\partial w_{12}^{(2)}}(X_2) = 2 \cdot (0{,}5 - 5) \cdot 1 \cdot 1{,}3 = (-11{,}7)$

$w_{12}^{(2)'} = 0 - 0{,}1 \cdot 0{,}5 \cdot (-0{,}5 + (-11{,}7)) = \boxed{0{,}61}$