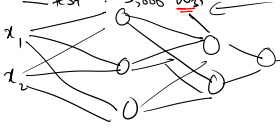
Dataset 10,000 — train : 7,000 — test : 3,000

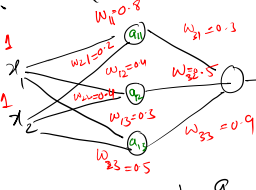


$J = \frac{1}{m}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$

$w_{i,new} = w_{i,old} - \alpha \frac{\partial J}{\partial w_i}$

Example:

backpropagation (backward pass)



$y = 2.2$

$J = (\hat{y} - y)^2$

$a_{11} = w_{11}x_1 + w_{21}x_2$

$= 0.8 \times 1 + 0.2 \times 1$

$a_{11} = 1$

$a_{12} = w_{12}x_1 + w_{22}x_2$

$= 0.4 \times 1 + 0.4 \times 1$

$a_{12} = 0.8$

$a_{13} = w_{13}x_1 + w_{23}x_2$

$= 0.3 \times 1 + 0.5 \times 1$

$a_{13} = 0.8$

$\hat{y} = w_{31}a_{11} + w_{32}a_{12} + w_{33}a_{13}$

$= 0.3 \times 1 + 0.5 \times 0.8 + 0.9 \times 0.8$

$= 0.3 + 0.4 + 0.72$

$\hat{y} = 1.42$

$w_{31,new} = w_{31,old} - \alpha \frac{\partial J}{\partial w_{31}}$, $\alpha = 0.01$

$= 0.3 - 0.01 \times -0.156$

$= 0.3 + 0.0156$

$w_{31,new} = 0.3156$

$$\frac{\partial J}{\partial w_{31}} = \underbrace{\frac{\partial J}{\partial \hat{y}}}_{1} \times \underbrace{\frac{\partial \hat{y}}{\partial w_{31}}}_{2} \quad \text{--- (A)}$$

① $$\frac{\partial J}{\partial \hat{y}} = \frac{\partial}{\partial \hat{y}} (\hat{y} - y)^2$$

$$= \frac{\partial}{\partial \hat{y}} (\hat{y}^2 + y^2 - 2\hat{y}y)$$

$$= 2\hat{y} - 2y$$

$$\frac{\partial J}{\partial \hat{y}} = 2(\hat{y} - y) \quad \text{--- (B)}$$

② $$\frac{\partial \hat{y}}{\partial w_{31}} = \frac{\partial}{\partial w_{31}} (w_{31} a_{11} + w_{32} a_{12} + w_{33} a_{13})$$

$$\frac{\partial \hat{y}}{\partial w_{31}} = a_{11} \rightarrow \text{(C)}$$

put (B) & (C) in (A)

$$\frac{\partial J}{\partial w_{31}} = 2(\hat{y} - y) \times a_{11}$$

$$= 2(1.42 - 2.2) \times 1$$

$$= -0.78$$

$$\frac{\partial J}{\partial w_{31}} = -0.156$$

$$\frac{\partial J}{\partial w_{11}} = \underbrace{\frac{\partial J}{\partial \hat{y}}}_{D} \times \underbrace{\frac{\partial \hat{y}}{\partial a_{11}}}_{E} \times \frac{\partial a_{11}}{\partial w_{11}} \rightarrow \text{(F)}$$

(D) $$\frac{\partial \hat{y}}{\partial a_{11}} = \frac{\partial}{\partial a_{11}} (w_{31} a_{11} + w_{32} a_{12} + w_{33} a_{13})$$

$$\frac{\partial \hat{y}}{\partial a_{11}} = w_{31} \rightarrow \text{(D)}$$

(E) $$\frac{\partial a_{11}}{\partial w_{11}} = \frac{\partial}{\partial w_{11}} (w_{11} x_1 + w_{21} x_2)$$

$$\frac{\partial a_{11}}{\partial w_{11}} = x_1 \rightarrow \text{(E)}$$

put (B), (D), (E) in (F)

$$\frac{\partial J}{\partial w_{11}} = 2(\hat{y} - y) \times w_{31} \times x_1$$

$$= -0.156 \times 0.3 \times 1$$

$$\frac{\partial J}{\partial w_{11}} = -0.0468$$

$$w_{11,new} = w_{11,old} - \alpha \frac{\partial J}{\partial w_{11}}$$

$$= 0.8 - 0.01 \times -0.0468$$

$$= 0.8 + 0.000468$$

$$\boxed{w_{11,new} = 0.800468}$$