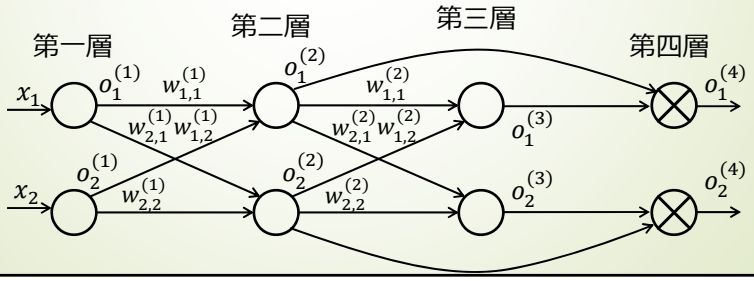
1

演習課題１２

- 下の多層ニューラルネットワークを考える．ただし誤差関数 $E = E(o_1^{(4)}, o_2^{(4)})$とし，$o_j^{(4)} = o_j^{(3)} * o_j^{(2)}$，$net_j^{(l)} = \sum_{i=1}^{2} o_i^{(l-1)} w_{j,i}^{(l-1)}$ $(l = 2,3)$，$o_j^{(l)} = f\left(net_j^{(l)}\right)$ $(l = 2,3)$ とする．このとき $\frac{\partial E}{\partial w_{1,1}^{(1)}}$ を求めよ．なお，閾値 θ は考えない．



12/1/2023

2

演習課題１２解答

- $\frac{\partial E}{\partial w_{1,1}^{(1)}} = \frac{\partial E}{\partial net_1^{(2)}} \boxed{\frac{\partial net_1^{(2)}}{\partial w_{1,1}^{(1)}}} = o_1^{(1)}$
- $\frac{\partial E}{\partial net_1^{(2)}} = \frac{\partial E}{\partial o_1^{(2)}} \boxed{\frac{\partial o_1^{(2)}}{\partial net_1^{(2)}}} = f'\left(net_1^{(2)}\right)$
- $\frac{\partial E}{\partial o_1^{(2)}} = \sum_j \frac{\partial E}{\partial o_j^{(4)}} \frac{\partial o_j^{(4)}}{\partial o_1^{(2)}} = \sum_j \frac{\partial E}{\partial o_j^{(4)}} \left(\frac{\partial o_j^{(3)}}{\partial o_1^{(2)}} o_j^{(2)} + \frac{\partial o_j^{(2)}}{\partial o_1^{(2)}} o_j^{(3)} \right)$
  $= \sum_j \left(\frac{\partial E}{\partial o_j^{(4)}} \boxed{\frac{\partial o_j^{(3)}}{\partial o_1^{(2)}}} o_j^{(2)} \right) + \frac{\partial E}{\partial o_1^{(4)}} o_1^{(3)}$
- $\boxed{\frac{\partial o_j^{(3)}}{\partial o_1^{(2)}}} = \frac{\partial o_j^{(3)}}{\partial net_j^{(3)}} \frac{\partial net_j^{(3)}}{\partial o_1^{(2)}} = f'(net_j^{(3)}) w_{j,1}^{(2)}$
- $\frac{\partial E}{\partial w_{1,1}^{(1)}} = o_1^{(1)} f'\left(net_1^{(2)}\right) \left\{ \frac{\partial E}{\partial o_1^{(4)}} o_1^{(3)} + \sum_j \left(\frac{\partial E}{\partial o_j^{(4)}} o_j^{(2)} f'\left(net_j^{(3)}\right) w_{j,1}^{(2)} \right) \right\}$

12/1/2023

3

演習課題１２解答（２）

- $\frac{\partial E}{\partial w_{1,1}^{(1)}} = \sum_j \frac{\partial E}{\partial o_j^{(4)}} \frac{\partial o_j^{(4)}}{\partial w_{1,1}^{(1)}} = \frac{\partial E}{\partial o_1^{(4)}} \left(\frac{\partial o_1^{(3)}}{\partial w_{1,1}^{(1)}} o_1^{(2)} + \frac{\partial o_1^{(2)}}{\partial w_{1,1}^{(1)}} o_1^{(3)} \right) + \frac{\partial E}{\partial o_2^{(4)}} \left(\frac{\partial o_2^{(3)}}{\partial w_{1,1}^{(1)}} o_2^{(2)} + \frac{\partial o_2^{(2)}}{\partial w_{1,1}^{(1)}} o_2^{(3)} \right)$ （$\frac{\partial o_2^{(2)}}{\partial w_{1,1}^{(1)}}$ は 0）
- $= \sum_j \frac{\partial E}{\partial o_j^{(4)}} \frac{\partial o_j^{(3)}}{\partial w_{1,1}^{(1)}} o_j^{(2)} + \frac{\partial E}{\partial o_1^{(4)}} \frac{\partial o_1^{(2)}}{\partial w_{1,1}^{(1)}} o_1^{(3)} = \sum_j \frac{\partial E}{\partial o_j^{(4)}} o_j^{(2)} \frac{\partial o_j^{(3)}}{\partial w_{1,1}^{(1)}} + \frac{\partial E}{\partial o_1^{(4)}} o_1^{(3)} \frac{\partial o_1^{(2)}}{\partial w_{1,1}^{(1)}}$
- $\frac{\partial o_j^{(3)}}{\partial w_{1,1}^{(1)}} = \frac{\partial o_j^{(3)}}{\partial net_j^{(3)}} \frac{\partial net_j^{(3)}}{\partial o_1^{(2)}} \frac{\partial o_1^{(2)}}{\partial w_{1,1}^{(1)}} = f'\left(net_j^{(3)}\right) w_{j,1}^{(2)} \frac{\partial o_1^{(2)}}{\partial w_{1,1}^{(1)}}$
- $\frac{\partial o_1^{(2)}}{\partial w_{1,1}^{(1)}} = \frac{\partial o_1^{(2)}}{\partial net_1^{(2)}} \frac{\partial net_1^{(2)}}{\partial w_{1,1}^{(1)}} = f'\left(net_1^{(2)}\right) o_1^{(1)}$
- $\frac{\partial E}{\partial w_{1,1}^{(1)}} = o_1^{(1)} f'\left(net_1^{(2)}\right) \left\{ \frac{\partial E}{\partial o_1^{(4)}} o_1^{(3)} + \sum_j \left(\frac{\partial E}{\partial o_j^{(4)}} o_j^{(2)} f'\left(net_j^{(3)}\right) w_{j,1}^{(2)} \right) \right\}$

12/1/2023