1

演習課題 6

- 誤差関数 $E = -\sum_{j=1}^{n(L)} t_j \ln\ o_j^{(L)}$ で与えられ，
出力層の活性化関数がソフトマックス関数
$f\left(net_j^{(L)}\right) = e^{net_j^{(L)}} / \sum_{m=1}^{n(L)} e^{net_m^{(L)}}$ で与えられた場合，
E を $w_{j,i}^{(L-1)}$ で偏微分した値が次式になることを示せ．
ただし，$\sum_{j=1}^{n(L)} t_j = 1$とする．

$$\frac{\partial E}{\partial w_{j,i}^{(L-1)}} = (o_j^{(L)} - t_j) o_i^{(L-1)}$$

10/19/2023

2

演習課題 6 解答

- $$\frac{\partial E}{\partial w_{j,i}^{(L-1)}} = \boxed{\frac{\partial E}{\partial net_j^{(L)}}} \boxed{\frac{\partial net_j^{(L)}}{\partial w_{j,i}^{(L-1)}}}$$

$$\frac{\partial net_j^{(L)}}{\partial w_{j,i}^{(L-1)}} = o_i^{(L-1)}$$

$$\boxed{\frac{\partial E}{\partial net_j^{(L)}}} = \sum_{k=1}^{n(L)} \frac{\partial E}{\partial o_k^{(L)}} \frac{\partial o_k^{(L)}}{\partial net_j^{(L)}}$$

$$= \sum_{k=1}^{n(L)} \left(-\frac{t_k}{o_k^{(L)}}\right) \boxed{\frac{\partial o_k^{(L)}}{\partial net_j^{(L)}}}$$

$net_j^{(L)} = \sum_{i=1}^{I+1} o_i^{(L-1)} w_{j,i}^{(L-1)}$
$(o_{I+1}^{(L-1)} = 1,\ w_{j,I+1}^{(L-1)} = -\theta_j^{(L)})$

$o_k^{(L)} = e^{net_k^{(L)}} / \sum_{j=1}^{n(L)} e^{net_j^{(L)}}$

⇧

$o_k^{(L)}$ が全ての $net_j^{(L)}$ に依存する

$E = -\sum_{j=1}^{n(L)} t_j \ln\ o_j^{(L)}$

10/19/2023

3

演習課題６解答

$o_k^{(L)} = e^{net_k^{(L)}} / \sum_{j=1}^{n(L)} e^{net_j^{(L)}}$

- $k = j$ のとき

$$\boxed{\frac{\partial o_k^{(L)}}{\partial net_j^{(L)}}} = \frac{\partial o_j^{(L)}}{\partial net_j^{(L)}} = \frac{e^{net_j^{(L)}} \sum_{m=1}^{n(L)} e^{net_m^{(L)}} - e^{net_j^{(L)}} e^{net_j^{(L)}}}{\left(\sum_{m=1}^{n(L)} e^{net_m^{(L)}}\right)^2}$$

$$= \frac{e^{net_j^{(L)}}}{\sum_{m=1}^{n(L)} e^{net_m^{(L)}}} - \left(\frac{e^{net_j^{(L)}}}{\sum_{m=1}^{n(L)} e^{net_m^{(L)}}}\right)^2$$

$$= o_j^{(L)} - {o_j^{(L)}}^2$$

- $k \neq j$ のとき

$$\boxed{\frac{\partial o_k^{(L)}}{\partial net_j^{(L)}}} = \frac{-e^{net_k^{(L)}} e^{net_j^{(L)}}}{\left(\sum_{m=1}^{n(L)} e^{net_m^{(L)}}\right)^2} = -o_k^{(L)} o_j^{(L)}$$

4

演習課題６解答

- $$\boxed{\frac{\partial E}{\partial net_j^{(L)}}} = \sum_{k=1}^{n(L)} \left(-\frac{t_k}{o_k^{(L)}}\right) \boxed{\frac{\partial o_k^{(L)}}{\partial net_j^{(L)}}}$$

$$= \sum_{k=1}^{n(L)} \left(-\frac{t_k}{o_k^{(L)}}\right)\left(-o_k^{(L)} o_j^{(L)}\right) + \left(-\frac{t_j}{o_j^{(L)}}\right) o_j^{(L)}$$

$$= \sum_{k=1}^{n(L)} t_k o_j^{(L)} - t_j$$

$$= o_j^{(L)} \sum_{k=1}^{n(L)} t_k - t_j$$

$$= o_j^{(L)} - t_j$$

- $$\frac{\partial E}{\partial w_{j,i}^{(L-1)}} = \boxed{\frac{\partial E}{\partial net_j^{(L)}}} \underbrace{\boxed{\frac{\partial net_j^{(L)}}{\partial w_{j,i}^{(L-1)}}}}_{= o_i^{(L-1)}} = (o_j^{(L)} - t_j) o_i^{(L-1)}$$