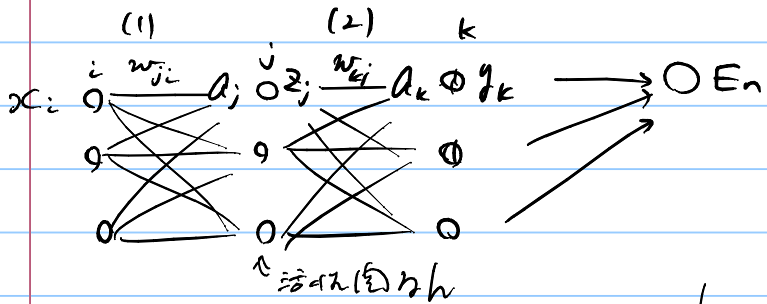
演習 5.22

ノートのタイトル

2014/07/06

ヘッセ行列の正確な値



$a_j = \sum_i w_{ji} x_i$, $z_j = h(a_j)$

$a_k = \sum_j w_{kj} z_j$, $y_k = g(a_k)$

$E_n = \frac{1}{2}\sum_k (y_k - t_k)^2$, $E = \sum_n E_n$

$\delta_k = \frac{\partial E_n}{\partial a_k}$, $M_{kk'} = \frac{\partial^2 E_n}{\partial a_k \partial a_{k'}}$ とおく

$a_k = \sum_j w_{kj} h(a_j)$ $\qquad \frac{\partial a_j}{\partial w_{ji}} = x_i$

$\therefore \frac{\partial a_k}{\partial a_j} = w_{kj} h'(a_j)$ $\qquad \frac{\partial a_k}{\partial w_{kj}} = z_j$

(5.93) $\frac{\partial E_n}{\partial w_{k'j'}} = \frac{\partial a_{k'}}{\partial w_{k'j'}}\frac{\partial E_n}{\partial a_{k'}} = z_{j'}\frac{\partial E_n}{\partial a_{k'}}$

$$\frac{\partial^2 E_n}{\partial w_{kj}\partial w_{k'j'}} = \frac{\partial a_k}{\partial w_{kj}}\frac{\partial}{\partial a_k}\left(z_{j'}\frac{\partial E_n}{\partial a_{k'}}\right)$$

$$= z_j z_{j'}\frac{\partial}{\partial a_k}\frac{\partial E_n}{\partial a_{k'}} = z_j z_{j'} M_{kk'}$$

$$\frac{\partial a_k}{\partial w_{ji}} = \frac{\partial}{\partial w_{ji}}\sum_j w_{kj} h(a_j)$$

$$= \frac{\partial}{\partial w_{ji}} w_{kj} h(a_j) = w_{kj} x_i h'(a_j)$$

(5.94)

$$\frac{\partial E_n}{\partial w_{ji}} = \frac{\partial a_j}{\partial w_{ji}}\frac{\partial E_n}{\partial a_j} = x_i \sum_k \frac{\partial a_k}{\partial a_j}\frac{\partial E_n}{\partial a_k} = x_i \sum_k w_{kj} h'(a_j)\delta_k = x_i h'(a_j)\sum_k w_{kj}\delta_k$$

$$\therefore \frac{\partial E_n}{\partial w_{j'i'}} = x_{i'} h'(a_{j'})\sum_k w_{kj'}\delta_k$$

$$\frac{\partial^2 E_n}{\partial w_{ji}\partial w_{j'i'}} = x_{i'}\left\{\overbrace{\frac{\partial}{\partial w_{ji}} h'(a_{j'})}^{\text{①}}\sum_k w_{kj'}\delta_k + h'(a_{j'})\sum_k w_{kj'}\overbrace{\frac{\partial}{\partial w_{ji}}\delta_k}^{\text{②}}\right\}$$

$$\text{①} = \frac{\partial h'(a_{j'})}{\partial w_{ji}} = \frac{\partial a_{j'}}{\partial w_{ji}} h''(a_{j'}) = x_i I_{jj'} h''(a_{j'}) \leftarrow I_{jj'} = \begin{cases} 0 & (j \neq j') \\ 1 & (j = j') \end{cases}$$

$$\text{②} = \frac{\partial \delta_k}{\partial w_{ji}} = \frac{\partial}{\partial w_{ji}}\frac{\partial E_n}{\partial a_k} = \sum_{k'}\frac{\partial a_{k'}}{\partial w_{ji}}\frac{\partial}{\partial a_{k'}}\frac{\partial E_n}{\partial a_k} = \sum_{k'}\frac{\partial a_{k'}}{\partial w_{ji}} M_{kk'}$$

$$= \sum_{k'} w_{k'j} x_i h'(a_j) M_{kk'} = x_i h'(a_j)\sum_{k'} w_{k'j} M_{kk'}$$

$$\therefore \frac{\partial^2 E_n}{\partial w_{ji} \partial w_{j'i'}} = x_{i'} x_i I_{jj'} h''(a_j) \sum_k w_{kj'} \delta_k + x_{i'} x_i h'(a_j) h'(a_j) \sum_k \sum_{k'} w_{kj'} w_{kj} M_{kk'}$$

$$(5.95) \quad \frac{\partial E_n}{\partial w_{kj'}} = \frac{\partial a_k}{\partial w_{kj'}} \frac{\partial E_n}{\partial a_k} = z_{j'} \delta_k$$

$$\frac{\partial^2 E_n}{\partial w_{ji} \partial w_{kj'}} = \frac{\partial z_{j'}}{\partial w_{ji}} \delta_k + z_{j'} \frac{\partial}{\partial w_{ji}} \delta_k$$

$$\frac{\partial z_{j'}}{\partial w_{ji}} = \frac{\partial}{\partial w_{ji}} h(a_{j'}) = \frac{\partial a_{j'}}{\partial w_{ji}} h'(a_{j'}) = x_i I_{jj'} h'(a_{j'})$$

$$\frac{\partial}{\partial w_{ji}} \delta_k = \frac{\partial}{\partial w_{ji}} \frac{\partial E_n}{\partial a_k} = \sum_{k'} \frac{\partial a_{k'}}{\partial w_{ji}} \frac{\partial^2 E_n}{\partial a_{k'} \partial a_k} = \sum_{k'} w_{k'j} x_i h'(a_j) M_{kk'}$$

$$\therefore \frac{\partial^2 E_n}{\partial w_{ji} \partial w_{kj'}} = x_i I_{jj'} h'(a_{j'}) \delta_k + z_{j'} \sum_{k'} w_{k'j} x_i h'(a_j) M_{kk'}$$

$$= x_i \left\{ I_{jj'} h'(a_{j'}) \delta_k + z_{j'} h'(a_j) \sum_{k'} w_{k'j} M_{kk'} \right\} \cdots$$

こっちが正しい
本文の(5.95)は
間違っている。

$$\frac{\partial}{\partial w_{ji}} \frac{\partial E_n}{\partial a_{k'}} = \sum_k \frac{\partial a_k}{\partial w_{ji}} \frac{\partial^2 E_n}{\partial a_k \partial a_{k'}} = \sum_k w_{kj} x_i h'(a_j) M_{kk'}$$

$$\therefore \frac{\partial^2 E_n}{\partial w_{ji} \partial w_{kj'}} = x_i I_{jj'} h'(a_{j'}) \frac{\partial E_n}{\partial a_k} + z_{j'} \sum_k w_{kj} x_i h'(a_j) M_{kk'}$$