演習5.16

y_n, t_nがベクトルのとき

$$E=\frac{1}{2}\sum_n\|y_n-t_n\|^2=\frac{1}{2}\sum_n\sum_i(y_{ni}-t_{ni})^2$$

$$\nabla E=\frac{1}{2}\sum_n\sum_i\nabla(y_{ni}-t_{ni})^2=\sum_n\sum_i(y_{ni}-t_{ni})\nabla y_{ni}$$

$$\nabla\nabla E=\sum_n\sum_i\nabla\{(y_{ni}-t_{ni})\nabla y_{ni}\}$$

$$=\sum_n\sum_i\nabla y_{ni}(\nabla y_{ni})^T+\sum_n\sum_i(y_{ni}-t_{ni})\nabla\nabla y_{ni}$$

第2項を無視すると(本文P.252の理由により)

$$\nabla\nabla E=\sum_n\sum_i\nabla y_{ni}(\nabla y_{ni})^T=\sum_n B_nB_n^T,\quad B_n=\nabla y_n$$

(最ゴの $\sum_n BB^T$ への書き換えについて)

① wが2次元, y_nが3次元の場合で様子を見る

$$\nabla y_n=\begin{pmatrix}\partial w_1\\ \partial w_2\end{pmatrix}(y_{n1}\ y_{n2}\ y_{n3})=\begin{pmatrix}\partial_{w_1}y_{n1} & \partial_{w_1}y_{n2} & \partial_{w_1}y_{n3}\\ \partial_{w_2}y_{n1} & \partial_{w_2}y_{n2} & \partial_{w_2}y_{n3}\end{pmatrix}$$

$$(\nabla y_n)^T=\begin{pmatrix}\partial_{w_1}y_{n1} & \partial_{w_2}y_{n1}\\ \partial_{w_1}y_{n2} & \partial_{w_2}y_{n2}\\ \partial_{w_1}y_{n3} & \partial_{w_2}y_{n3}\end{pmatrix}$$

$$\therefore\ \nabla y_n(\nabla y_n)^T=\begin{pmatrix}(\partial_{w_1}y_{n1})^2+(\partial_{w_1}y_{n2})^2+(\partial_{w_1}y_{n3})^2 & \partial_{w_1}y_{n1}\partial_{w_2}y_{n1}+\partial_{w_1}y_{n2}\partial_{w_2}y_{n2}+\partial_{w_1}y_{n3}\partial_{w_2}y_{n3}\\ \partial_{w_2}y_{n1}\partial_{w_1}y_{n1}+\partial_{w_2}y_{n2}\partial_{w_1}y_{n2}+\partial_{w_2}y_{n3}\partial_{w_1}y_{n3} & (\partial_{w_2}y_{n1})^2+(\partial_{w_2}y_{n2})^2+(\partial_{w_2}y_{n3})^2\end{pmatrix}$$

$$\nabla y_{n1}=\begin{pmatrix}\partial_{w_1}y_{n1}\\ \partial_{w_2}y_{n1}\end{pmatrix}$$

$$\therefore\ \nabla y_{n1}(\nabla y_{n1})^T=\begin{pmatrix}(\partial_{w_1}y_{n1})^2 & \partial_{w_1}y_{n1}\partial_{w_2}y_{n1}\\ \partial_{w_1}y_{n1}\partial_{w_2}y_{n1} & (\partial_{w_2}y_{n1})^2\end{pmatrix}$$

$$\nabla y_{n1}(\nabla y_{n1})^T+\nabla y_{n2}(\nabla y_{n2})^T+\nabla y_{n3}(\nabla y_{n3})^T=\nabla y_n(\nabla y_n)^T$$

$$\therefore\ \nabla y_n(\nabla y_n)^T=\sum_i\nabla y_{ni}(\nabla y_{ni})^T$$

② block matrixの積を使ってもう少し見通し良くする。

$$\nabla y_n = \begin{pmatrix} \partial w_1 \\ \partial w_2 \end{pmatrix} \begin{pmatrix} y_{n1} & y_{n2} & y_{n3} \end{pmatrix} = \left(\begin{pmatrix} \partial w_1 \\ \partial w_2 \end{pmatrix} y_{n1} \quad \begin{pmatrix} \partial w_1 \\ \partial w_2 \end{pmatrix} y_{n2} \quad \begin{pmatrix} \partial w_1 \\ \partial w_2 \end{pmatrix} y_{n3} \right)$$

$$= \begin{pmatrix} \nabla y_{n1} & \nabla y_{n2} & \nabla y_{n3} \end{pmatrix}$$

$$(\nabla y_n)^T = \begin{pmatrix} (\partial w_1 \ \partial w_2)\, y_{n1} \\ (\partial w_1 \ \partial w_2)\, y_{n2} \\ (\partial w_1 \ \partial w_2)\, y_{n3} \end{pmatrix} = \begin{pmatrix} (\nabla y_{n1})^T \\ (\nabla y_{n2})^T \\ (\nabla y_{n3})^T \end{pmatrix}$$

$$\nabla y_n (\nabla y_n)^T = \begin{pmatrix} \nabla y_{n1} & \nabla y_{n2} & \nabla y_{n3} \end{pmatrix} \begin{pmatrix} (\nabla y_{n1})^T \\ (\nabla y_{n2})^T \\ (\nabla y_{n3})^T \end{pmatrix}$$

← block matrixの積

$$= \nabla y_{n1} (\nabla y_{n1})^T + \nabla y_{n2} (\nabla y_{n2})^T + \nabla y_{n3} (\nabla y_{n3})^T$$

$$= \sum_i \nabla y_{ni} (\nabla y_{ni})^T$$