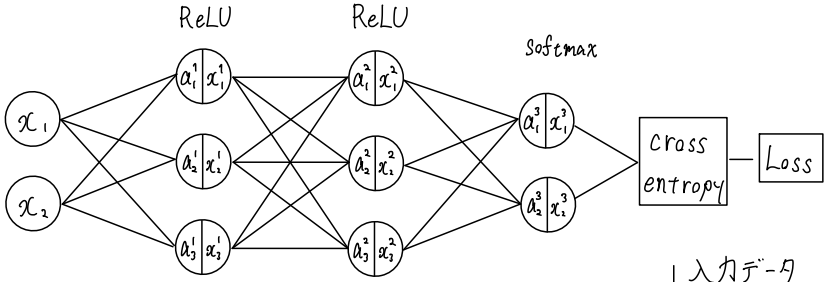

$$W^1 = \begin{pmatrix} 1 & -2 & 3 \\ -1 & 2 & -3 \end{pmatrix},\ W^2 = \begin{pmatrix} -1 & 2 & -3 \\ 1 & -2 & 3 \\ -1 & 2 & -3 \end{pmatrix},\ W^3 = \begin{pmatrix} -1 & 2 \\ -3 & 1 \\ -2 & 3 \end{pmatrix}$$

$$B^1 = \begin{pmatrix} 1 & -2 & 3 \end{pmatrix},\ B^2 = \begin{pmatrix} 1 & -2 & 3 \end{pmatrix},\ B^3 = \begin{pmatrix} 34 & -54 \end{pmatrix}$$

入力データ
$X = (-1\ 2)$
教師ラベル
$T = (1\ 0)$
学習率
$\eta = 0.1$

(1) $A^1 = XW^1 + B^1 = (-1\ 2)\begin{pmatrix} 1 & -2 & 3 \\ -1 & 2 & -3 \end{pmatrix} + (1\ -2\ 3)$

$= (-3\ \ 6\ \ -9) + (1\ -2\ 3)$

$= (-2\ \ 4\ \ -6)$

(2) $X^1 = ReLU(A^1) = (0\ \ 4\ \ 0)$

(3) $A^2 = X^1W^2 + B^2 = (0\ \ 4\ \ 0)\begin{pmatrix} -1 & 2 & -3 \\ 1 & -2 & 3 \\ -1 & 2 & -3 \end{pmatrix} + (1\ -2\ 3)$

$= (4\ -8\ 12) + (1\ -2\ 3)$

$= (5\ -10\ 15)$

(4) $X^2 = ReLU(A^2) = (5\ \ 0\ \ 15)$

(5) $A^3 = X^2W^3 + B^3 = (5\ \ 0\ \ 15)\begin{pmatrix} -1 & 2 \\ -3 & 1 \\ -2 & 3 \end{pmatrix} + (34\ -54)$

$= (-35\ \ 55) + (34\ -54)$

$= (-1\ \ 1)$

(6) $X^3 = \left(\frac{e^{-1}}{e^{-1}+e^{1}}\ \ \frac{e^{1}}{e^{-1}+e^{1}}\right)$

$= \left(\frac{e^{-2}}{e^{-2}+1}\ \ \frac{1}{e^{-2}+1}\right)$ $\quad e^{-2} = 0.135$

$= \left(\frac{0.135}{1.135}\ \ \frac{1}{1.135}\right)$

$= (0.119\ \ 0.881)$

(7) $CE(X^3, T)$

$= -1 \cdot \ln 0.119 - 0 \cdot \ln 0.881$

$= 2.129$

(8) $\frac{\partial L}{\partial X^3} = \left(\frac{\partial L}{\partial x_1^3}\ \ \frac{\partial L}{\partial x_2^3}\right) = (-8.4\ \ 8.4)$

$\frac{\partial L}{\partial x_1^3} = -\frac{t_1}{x_1^3} + \frac{t_2}{x_2^3} = -\frac{1}{0.119} = -8.40$

$\frac{\partial L}{\partial x_2^3} = -\frac{t_2}{x_2^3} + \frac{t_1}{x_1^3} = \frac{1}{0.119} = 8.40$

(9) $\frac{\partial h_3}{\partial A^3} = \left(\frac{\partial h_3}{\partial a_1^3}\ \ \frac{\partial h_3}{\partial a_2^3}\right) = (0.1048\ \ 0.1048)$

$\frac{\partial h_3}{\partial a_1^3} = x_1^3(1-x_1^3) = 0.119 \times 0.881$

$= x_2^3(1-x_2^3)$

(10) $\Delta^3 = \frac{\partial L}{\partial X^3} \odot \frac{\partial h_3}{\partial A^3} = (-0.88\ \ 0.88)$

(11) $\frac{\partial L}{\partial B^3} = \Delta^3 = (-0.88\ \ 0.88)$

(12) $\frac{\partial L}{\partial W^3} = X^{2T}\Delta^3 = \begin{pmatrix} 5 \\ 0 \\ 15 \end{pmatrix}(-0.88\ \ 0.88)$

$= \begin{pmatrix} -4.4 & 4.4 \\ 0 & 0 \\ -13.2 & 13.2 \end{pmatrix}$

(13) $\frac{\partial h_2}{\partial A^2} = \begin{pmatrix} \frac{\partial h_2}{\partial a_1^2} & \frac{\partial h_2}{\partial a_2^2} & \frac{\partial h_2}{\partial a_3^2} \end{pmatrix}$

$= \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}$

(14) $\Delta^2 = \Delta^3 W^{3T} \odot \frac{\partial h_2}{\partial A^2}$

$= \begin{pmatrix} -0.88 & 0.88 \end{pmatrix} \begin{pmatrix} -1 & -3 & -2 \\ 2 & 1 & 3 \end{pmatrix} \odot \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}$

$= \begin{pmatrix} 2.64 & 3.52 & 4.4 \end{pmatrix} \odot \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}$

$= \begin{pmatrix} 2.64 & 0 & 4.4 \end{pmatrix}$

(15) $\frac{\partial L}{\partial B^2} = \Delta^2 = \begin{pmatrix} 2.64 & 0 & 4.4 \end{pmatrix}$

(16) $\frac{\partial L}{\partial W^2} = X^{1T} \Delta^2 = \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix} \begin{pmatrix} 2.64 & 0 & 4.4 \end{pmatrix}$

$= \begin{pmatrix} 0 & 0 & 0 \\ 10.56 & 0 & 17.6 \\ 0 & 0 & 0 \end{pmatrix}$

(17) $\frac{\partial h_1}{\partial A^1} = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$

(18) $\Delta^1 = \Delta^2 W^{2T} \odot \frac{\partial h_1}{\partial A^1}$

$= \begin{pmatrix} 2.64 & 0 & 4.4 \end{pmatrix} \begin{pmatrix} -1 & 1 & 1 \\ 2 & -2 & 2 \\ -3 & 3 & -3 \end{pmatrix} \odot \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$

$= \begin{pmatrix} -15.84 & 15.84 & -10.56 \end{pmatrix} \odot \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$

$= \begin{pmatrix} 0 & 15.84 & 0 \end{pmatrix}$

(19) $\frac{\partial L}{\partial B^1} = \Delta^1 = \begin{pmatrix} 0 & 15.84 & 0 \end{pmatrix}$

(20) $\frac{\partial L}{\partial W^1} = X^{0T} \Delta^1 = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \begin{pmatrix} 0 & 15.84 & 0 \end{pmatrix}$

$= \begin{pmatrix} 0 & -15.84 & 0 \\ 0 & 31.68 & 0 \end{pmatrix}$

(21) $W^1 = W^1 - \eta \frac{\partial L}{\partial W^1}$

$= \begin{pmatrix} 1 & -2 & 3 \\ -1 & 2 & -3 \end{pmatrix} - 0.1 \begin{pmatrix} 0 & -15.84 & 0 \\ 0 & 31.68 & 0 \end{pmatrix}$

$= \begin{pmatrix} 1 & -0.416 & 3 \\ -1 & -1.168 & -3 \end{pmatrix}$

(22) $W^2 = W^2 - \eta \frac{\partial L}{\partial W^2}$

$= \begin{pmatrix} -1 & 2 & -3 \\ 1 & -2 & 3 \\ -1 & 2 & -3 \end{pmatrix} - 0.1 \begin{pmatrix} 0 & 0 & 0 \\ 10.56 & 0 & 17.6 \\ 0 & 0 & 0 \end{pmatrix}$

$= \begin{pmatrix} -1 & 2 & -3 \\ -0.056 & -2 & 1.24 \\ -1 & 2 & -3 \end{pmatrix}$

(23) $W^3 = W^3 - \eta \frac{\partial L}{\partial W^3}$

$= \begin{pmatrix} -1 & 2 \\ -3 & 1 \\ -2 & 3 \end{pmatrix} - 0.1 \begin{pmatrix} -4.4 & 4.4 \\ 0 & 0 \\ -13.2 & 13.2 \end{pmatrix}$

$= \begin{pmatrix} -0.56 & 1.56 \\ -3 & 1 \\ -0.68 & 1.68 \end{pmatrix}$

(24) $B^1 = B^1 - \eta \frac{\partial L}{\partial B^1}$

$= \begin{pmatrix} 1 & -2 & 3 \end{pmatrix} - 0.1 \begin{pmatrix} 0 & 15.84 & 0 \end{pmatrix}$

$= \begin{pmatrix} 1 & -3.58 & 3 \end{pmatrix}$

(25) $B^2 = B^2 - \eta \frac{\partial L}{\partial B^2}$

$= \begin{pmatrix} 1 & -2 & 3 \end{pmatrix} - 0.1 \begin{pmatrix} 2.64 & 0 & 4.4 \end{pmatrix}$

$= \begin{pmatrix} 0.736 & -2 & 2.56 \end{pmatrix}$

(26) $B^3 = B^3 - \eta \frac{\partial L}{\partial B^3}$

$= \begin{pmatrix} 34 & -54 \end{pmatrix} - 0.1 \begin{pmatrix} -0.88 & 0.88 \end{pmatrix}$

$= \begin{pmatrix} 34.1 & -54.1 \end{pmatrix}$

(27) $CE(X^3, T) = 0$