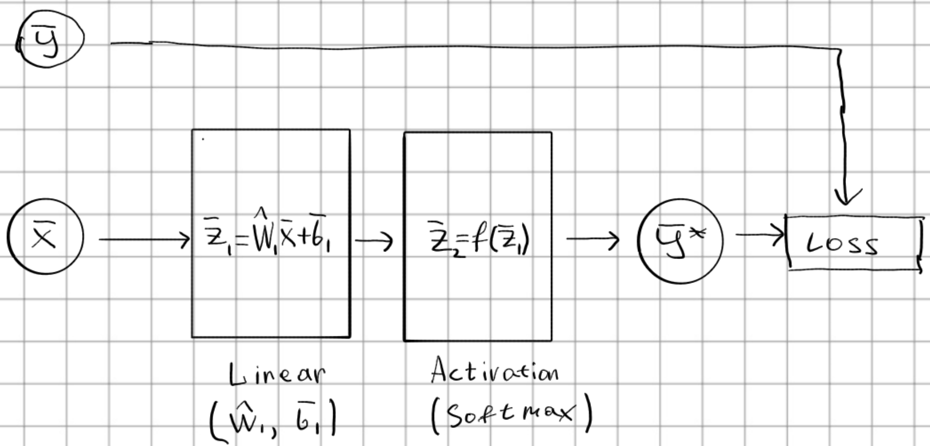

пр. $\bar{x} = \begin{pmatrix} 0.3 \\ 0.3 \\ 0.9 \end{pmatrix}$ $\qquad \bar{y} = \begin{pmatrix} 0.0 \\ 1.0 \end{pmatrix}$ $\begin{matrix}\text{ехать} \\ \text{стоять}\end{matrix}$

$$\bar{W}_1 = \begin{pmatrix} 0.2 & 0.4 & 0.9 \\ 0.1 & -0.2 & -0.5 \end{pmatrix} \qquad \bar{b}_1 = \begin{pmatrix} 0.0 \\ 0.2 \end{pmatrix}$$

$$\text{Softmax}(\bar{z}) = \frac{e^{z^i}}{\sum_k e^{z^k}}, \qquad \text{Loss} = (\bar{y} - \bar{y}^*)^2$$

I. Forward

1) $\bar{z}_1 = \bar{W}_1\bar{x} + \bar{b}_1 = \begin{pmatrix} 0.2 & 0.4 & 0.9 \\ 0.1 & -0.2 & -0.5 \end{pmatrix} \times \begin{pmatrix} 0.3 \\ 0.3 \\ 0.9 \end{pmatrix} + \bar{b}$

$$= \begin{pmatrix} 0.99 \\ -0.48 \end{pmatrix} + \begin{pmatrix} 0.0 \\ 0.2 \end{pmatrix} = \begin{pmatrix} 0.99 \\ -0.28 \end{pmatrix}$$

2) $\bar{z}_2 = \text{Softmax}(\bar{z}_1)$ $\qquad e^{\bar{z}_1} / \sum e^{\bar{z}_1}$

$$e^{\bar{z}_1} = \exp\begin{pmatrix} 0.99 \\ -0.28 \end{pmatrix} \simeq \begin{pmatrix} 2.71 \\ 0.7 \end{pmatrix}$$

$\sum = 3.41$

$$\bar{z}_2 = \begin{pmatrix} 0.8 \\ 0.2 \end{pmatrix}$$

3) $\bar{y}^* = \bar{z}$

Loss: $\cdot\ \Delta\bar{y} = \bar{y} - \bar{y}^* = \begin{pmatrix} 0. \\ \end{pmatrix} - \begin{pmatrix} 0.8 \\ \end{pmatrix} = \begin{pmatrix} -0.8 \\ \end{pmatrix}$

$\frac{\partial}{\partial x} = \partial_x$

$\begin{pmatrix} 1 \end{pmatrix} \quad \begin{pmatrix} 0.2 \end{pmatrix} \quad \begin{pmatrix} 0.8 \end{pmatrix}$

- $(\Delta \bar{y})^2 \simeq 2^7 \cdot 0.1^2 = 1.28$

$\partial_x f(g(x))$
$= \partial_g f \cdot \partial_x g$

II. Backward

$\partial_{Loss} Loss = 1 \qquad \partial_x x = 1$

$A(x) = y - x$

$(A)^2$

$\partial_x (A)^2 = 0$

$\partial_{\bar{y}^*} Loss : \quad \partial_{\bar{y}^*} (\bar{y} - \bar{y}^*)^2 = 2(\bar{y} - \bar{y}^*) \cdot \partial_{\bar{y}^*} (\bar{y} - \bar{y}^*) =$

$= -2(\bar{y} - \bar{y}^*) = \begin{pmatrix} 1.6 \\ -1.6 \end{pmatrix}$

3) $\bar{y}^* = \bar{z}_2 \qquad \underline{\partial_{\bar{z}_2} Loss} = \partial_{\bar{z}_2} \bar{y}^* \, \underline{\partial_{\bar{y}^*} Loss}$

$\underline{\overset{*}{y}{}^i = z_2^i}$

$\frac{\partial}{\partial z_2^j} \overset{*}{y}{}^i = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \hat{1} \begin{pmatrix} 1.6 \\ -1.6 \end{pmatrix} = \begin{pmatrix} 1.6 \\ -1.6 \end{pmatrix}$

2) $\partial_{\bar{z}_1} Loss = \partial_{\bar{z}_1} \bar{z}_2 \; \partial_{\bar{z}_2} Loss$

- $\partial_{\bar{z}_1} \bar{z}_2 = \partial_{\bar{z}_1} Softmax(\bar{z}_1), \quad Softmax(\bar{x})_i = \frac{e^{x_i}}{\sum_k e^{x_k}}$

$\underline{\partial_{x^j} Softmax(x^i)} = \begin{cases} j = i : & softmax(x^i)\,(1 - softmax(x^i)) \\ j \neq i : & -\, softmax(x^i)\, softmax(x^j) \end{cases}$

- $z_2^i = Softmax(z_1^i) = \begin{pmatrix} 0.8 \\ 0.2 \end{pmatrix}$

$\frac{\partial}{\partial z_1^j} z_2^i = \begin{cases} j = i : & z_2^i (1 - z_2^i) \\ j \neq i : & -z_2^i z_2^j \end{cases} = \begin{array}{c|c} & j \\ \hline i & \begin{pmatrix} 0.8(1-0.8) & -0.8 \cdot 0.2 \\ -0.2 \cdot 0.8 & 0.2(1-0.2) \end{pmatrix} \end{array}$

$= 0.16 \cdot \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

$\partial_{\bar{z}_1} Loss = 0.16 \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1.6 \\ -1.6 \end{pmatrix} = 0.16 \cdot \underset{2^9 \cdot 10^{-3}}{3.2} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

---

3) $\cdot \; \partial_{\bar{b}} Loss = \partial_{\bar{b}} \bar{z}_1 \cdot \partial_{\bar{z}_1} Loss$

$\partial_{\bar{b}} \bar{z}_1 = \partial_{\bar{b}} (\cancel{\hat{W} \bar{x}} + \bar{b}) = \hat{1}$

$$\partial_{\bar{b}} \text{Loss} = \partial_{\bar{z}_1} \text{loss} = 2^9 \cdot 10^{-3} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

- $$\partial_{\hat{W}} \text{Loss} = \partial_{\hat{W}} \bar{Z}_1 \cdot \partial_{\bar{z}_1} \text{Loss}$$

$$\partial_{\hat{w}} \bar{Z}_1 = \partial_{\hat{w}} (\hat{W} \bar{x} + \cancel{\bar{b}}) = \bar{x} = \begin{pmatrix} 0.3 \\ 0.3 \\ 0.9 \end{pmatrix}$$

$$\partial_{\hat{w}} \text{loss} = \begin{pmatrix} 0.3 \\ 0.3 \\ 0.9 \end{pmatrix} \times \begin{pmatrix} 1 \\ -1 \end{pmatrix} \cdot 2^9 \cdot 10^{-3}$$

$$= \begin{pmatrix} 0.3 & 0.3 & 0.9 \\ -0.3 & -0.3 & -0.9 \end{pmatrix} \cdot 2^9 \cdot 10^{-3}$$

$$\frac{2^{10} \cdot 10^{-3}}{2} \simeq 0.5$$

$$= \begin{pmatrix} 0.15 & 0.15 & 0.45 \\ -0.15 & -0.15 & -0.45 \end{pmatrix}$$