Backpropagation

Tuesday, April 23, 2024 11:04 AM

Forward Propagation

$$\underset{\text{Input}}{X \in \mathbb{R}^D} \rightsquigarrow \overset{Z_1}{W_1 X + b_1} \rightsquigarrow \overset{A_1}{\sigma(Z_1)}$$

$$Z_2 := W_2 A_1 + b_2$$

$$A_2 := \text{Softmax}(Z_2)$$

W_1 is $h \times D$ matrix

h is # of hidden nodes

$b_1, b_2 \in \mathbb{R}$

W_2 is $O \times h$ matrix

O is # of output nodes

$\sigma(x) = \frac{1}{1+e^{-x}}$ sigmoid function

$\text{Softmax}\left(\begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}\right) = \frac{1}{\sum e^{x_i}} \left(e^{x_1}, \dots, e^{x_k}\right)$

We use CROSS ENTROPY LOSS

$$L(y, \hat{y}) = -\sum y_i \log \hat{y}_i$$

BACKPROPAGATION

$$L(y, \hat{y}) = L(y, A_2) = L(y, \text{Softmax } Z_2) =$$

$$= -\sum y_i \log\left(\frac{e^{z_{2,i}}}{\sum_j e^{z_{2,j}}}\right) = -\sum y_i \left(z_{2,i} - \log\left(\sum_j e^{z_{2,j}}\right)\right) =$$

$$= \log\left(\sum_j e^{z_{2,j}}\right) - \sum y_i z_{2,i} \quad \left(\text{we used } \sum y_i = 1\right).$$

- $\frac{\partial L}{\partial Z_{2,i}} = \frac{e^{z_{2,i}}}{\sum_j e^{z_{2,j}}} - y_i = \hat{y}_i - y_i$. $\frac{\partial L}{\partial Z_2} = \hat{y} - y$

- $\frac{\partial L}{\partial A_1} = \frac{\partial L}{\partial Z_2} \cdot \frac{\partial Z_2}{\partial A_1} = \frac{\partial L}{\partial Z_2} \cdot W_2$

- $\frac{\partial L}{\partial b_2} = \frac{\partial L}{\partial Z_2} \cdot \frac{\partial Z_2}{\partial b_2} = \frac{\partial L}{\partial Z_2} \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} =$ sum of components of $\frac{\partial L}{\partial Z_2}$

- $\frac{\partial L}{\partial Z_1} = \frac{\partial L}{\partial A_1} \cdot \frac{\partial A_1}{\partial Z_1} = \frac{\partial L}{\partial A_1} \cdot$ (matrix Z_1, but apply $\frac{d\sigma}{dx}$ to every element)

$$\frac{d\sigma}{dx}(x) = \frac{e^{-x}}{(1+e^{-x})^2}$$

- $\frac{\partial L}{\partial A_1} = \frac{\partial L}{\partial Z_1} \cdot \frac{\partial Z_1}{\partial A_1} = \frac{\partial L}{\partial Z_1} \cdot W_1$
- $\frac{\partial L}{\partial b_1} = \frac{\partial L}{\partial Z_1} \cdot \frac{\partial Z_1}{\partial b_1} = \frac{\partial L}{\partial Z_1} \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} =$

= sum of elements of $\frac{\partial L}{\partial Z_1}$