Assignment 1

Monday, 23 April 2018 19:04

softmax.

$L_i = -f_{y_i} + \log \sum_j e^{f_j}$.

$\frac{dL_i}{dw_{y_i}} = -x_{y_i} + \frac{x_{y_i} e^{f_{y_i}}}{\sum_j e^{f_j}}$.

$\frac{\partial L_i}{\partial w_j} = \frac{x_j}{\sum_j e^{f_j}}$.

$L_i = -f_{y_i} + \log \sum e^{f_j}$

$= -x + \frac{e^{f_j} \cdot w_j}{\log \sum e^{f_j}}$

$\frac{\partial \ln(\sum e^{f_j})}{\partial y_i} \to \frac{e^{f_{y_i}}}{\sum e^{f_j}} \to y_i$.

$-y_i \frac{\partial \ln \sum}{\partial y_i}$

np.log.

$N \times C$

$D \times N$

$N \times D$.

$D \times N \times N \times L$

$x_1 =$

$N \times D$.

$dw_j = \text{np.sum}($

$N \times D \times C$

$D \times N \cdot N \times L$

$[x_1, x_2, x_3]$

$D \times L \cdot [$

$]$

$[x_1, x_2, x_3, x_4] \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [\overbrace{x_2, x_1, x_3, x_4}^{D \times C}]$.

X $N \times D$.

$q =$

W $D \times 1$

$N \times D \times D$

$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} [w_1 \cdot w_2, w_3 \ldots w_n]$

$N \times D$

$a = \begin{bmatrix} x_1 w_1, & x_1 w_2 \ldots & x_1 w_n \\ & & \\ x_n w_1, & \ldots & , x_n w_n \end{bmatrix}$

$\frac{dq}{dw_1} = \underline{x_1 \cdot w_1}$, $N \times D \times C$.

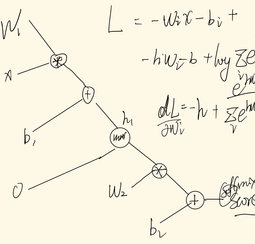
$x_1 \cdot w_1$

$q = x \cdot w$.

$[x_1, x_2] \cdot \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = [w_1 x_1 + w_3 x_2, w_2 x_1 + w_4 x_2]$.

$\frac{dq}{dw_1} = [$

$\frac{h w_2 + b}{\sum_j e^{h_j w_2 + b_2}}$.



$L = -w_i x - b_i +$

$-h w_i - b + \log \sum e^{h w_i}$

$\frac{e^{h w_i}}{\sum_j e^{h w_j}} \cdot h$.

$\frac{dL}{dw_i} = h + \sum_j e^{h w_j}$

$\lambda \cdot w^T w$.

$-\log \sum$

$\frac{e^{f_{y_i}} \cdot f_{y_i} \sum}{\sum_j e^{f_j}}$

$\frac{1}{N} \sum$

$\begin{bmatrix} x_{11} & x_{12} \end{bmatrix}$ each.

$D \times C$

b_i.

$1 \times C$

$N \times 1$

$\begin{bmatrix} \end{bmatrix}$

$w + w$

$h \cdot c$.

softmax(h)

$= \frac{e^{h w_i}}{\sum_j e^{h w_j}}$

$\frac{dS}{dh} = -w_y^T + \frac{e^{h w_i} \cdot w_i}{\sum_j e^{h w_j}}$

$= \frac{(e^{h w_i} - 1) w_y^T}{\sum_j e^{h w_j}}$

$\frac{dS}{dw} = -h^T + \frac{e^{h w_i} h^T}{\sum_j e^{h w_j}}$

h: $N \times 1$.

$dw_i = \frac{(e^{h w_i} - 1) h^T}{\sum_j e^{h w_j}}$

$N \times D$.

$\frac{ds}{db} = -h^T +$

$\frac{ds}{df} = N \times C$

$\frac{ds}{df} = -df + \frac{e^{f_i} \cdot df_i}{\sum_j e^{f_j}}$.

$N \times C$

$b \times n$.

$\frac{ds}{dw} = -X^T + \frac{e^{f_i} \cdot X^T}{\sum_j e^{f_j}} = X^T \frac{ds}{df}$.

$1 \times C$.

$\frac{ds}{dx} = -W^T + \frac{e^{f_i} \cdot W^T}{\sum_j e^{f_j}} = \frac{ds}{df} \cdot W^T$

$D \times 1 \times H \times L$

$\frac{d(x \cdot w)}{dW} = X^T dw$

$D \times 1$.

$N \times D$

$D \times N \times N \times 1$

$N \times 1$.

$\frac{d(x \cdot w)}{dx} = dW^T$.

$dh1$: $1 \times C$.

$h1 = \max(0, wX + b)$

$N \times D \times D \times C$

$N \times C$

X: N ** $N \times 1$.

W: $D \times 1$.

$h1$: $N \times 1$.

$wX + b$.

$N \times 1$.

max

$D \times N$

$0 \quad X^T (wX + b > 0) \times$

$D \times N \times N \times 1$.

$dh1$.

$h1 > 0$