Deep Learning Assignment1

Xiao Jing xj655

February 2019

1 Backpropagation

1. Apply the chain rule to:
$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial W}$$
While
$$\frac{\partial y}{\partial W} = x$$
And x is known, then
$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial y}x$$
Redo the process to $\frac{\partial L}{\partial b}$
$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial b}$$
While
$$\frac{\partial y}{\partial b} = 1$$
And x is known, then
$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial y}$$

2. Using the expression of y given in the question:
$$\frac{\partial y_j}{\partial x_i} = \frac{\partial}{\partial x_i}\Big(\frac{exp(\beta x_j)}{\sum_i exp(\beta x_i)}\Big)$$
In the case $i = j$:
$$\frac{\partial y_j}{\partial x_i} = \frac{\beta exp(\beta x_j)(\sum_i exp(\beta x_i)) - exp(\beta x_j)\beta exp(\beta x_j)}{(\sum_i exp(\beta x_i))^2} =$$
$$= \frac{\beta exp(\beta x_j)}{\sum_i exp(\beta x_i)}\Big[1 - \frac{exp(\beta x_j)}{\sum_i exp(\beta x_i)}\Big]$$

And substituting the expression of y_j:

$$\frac{\partial y_j}{\partial x_i} = \beta y_j(1 - y_j)$$

In the case $i \neq j$:

$$\frac{\partial y_j}{\partial x_i} = exp(\beta x_j)\frac{-\beta exp(\beta x_i)}{(\sum_i exp(\beta x_i))^2} =$$

$$= \beta \frac{exp(\beta x_j)}{\sum_i exp(\beta x_i)} \frac{exp(\beta x_i)}{\sum_i exp(\beta x_i)} = \beta y_j y_i$$

The answer for the question is: $\beta y_j(1 - y_j)\, if$ i= j
$\beta y_j y_i\, if$ i$\neq j$