Homework1

Xiong Jiao

October 24, 2019

1.

Assume the activation function is a linear function. i.e.

$$f(x) = ax + b$$

Then,

$$f\left(\Sigma\omega_{ji}x_i + w_{j0}\right) = a\sum_i w_{ji}x_i + b\sum_i w_{ji} + w_{j0}$$

$$g_k(x) = y_k = f\left(\sum_j w_{kj}\left(a\sum_i w_{ji}x_i + b\sum_j w_{ji} + w_{j0}\right) + w_{k0}\right)$$

$$= a^2\sum_j\sum_i w_{kj}w_{ji}x_i + ab\sum_j\sum_i w_{kj}w_{ji} + b\sum_i w_{ji} + w_{k0}$$

Finally, the result is only related to *x*, so it's linear.

2.

(1)

$$E = \frac{1}{2}(g-y)^2 = \frac{1}{2}\left(g - \frac{1}{1+e^{-w^T x}}\right)^2$$

$$\frac{\partial E}{\partial w} = \left(g - \frac{1}{1+e^{-w^T x}}\right)\cdot\frac{1}{\left(1+e^{-w^T x}\right)^2}\cdot e^{-w^T x}\cdot -x$$

$$= -\left(g - \frac{1}{1+e^{-w^T x}}\right)\cdot\frac{e^{-w^T x}x}{\left(1+e^{-w^T x}\right)^2}$$

The weight-updating formula equals to

$$w - \lambda\frac{\partial E}{\partial w}$$

$$= w + \lambda\left(g - \frac{1}{1+e^{-w^T x}}\right)\cdot\frac{e^{-w^T x}x}{\left(1+e^{-w^T x}\right)^2}$$

(2)

$$w=(0.5,1,1)^T, x=(1,2,0.5)^T, g=1, \lambda=0.1$$

$$w^T x=3$$

$$w_{new}=(0.5,1,1)^T+0.1(1-\frac{1}{1+e^{-3}})\cdot\frac{e^{-3}(1,2,0.5)^T}{\left(1+e^{-3}\right)^2}$$

$$=\begin{pmatrix}0.5\\1\\1\end{pmatrix}+2.1\times10^{-4}\begin{pmatrix}1\\2\\0.5\end{pmatrix}=\begin{pmatrix}0.5002\\1.0004\\1.0001\end{pmatrix}$$

3.

f_1 : the activation function.

f_2 : the function of output layer.

N_ℓ : the neuron number of the ℓth layer.

E : the error function.

Z : the value of samples.

λ : the learning rate.

Forward Pass

The first layer:

$$input_{j_1}^1=\sum_{i=1}^{n}w_{ij_1}^1 x_i$$

$$output_{j_1}^1=f_1(input_{j_1}^1)$$

The second layer:

$$input_{j_2}^2=\sum_{j_1=1}^{N_1}w_{j_1 j_2}^2 output_{j_1}^1$$

$$output_{j_2}^2=f_1(input_{j_2}^2)$$

The output layer:

$$input_{j_3}^3 = \sum_{j_2=1}^{N_2} w_{j_2 j_3}^3 output_{j_2}^2$$

$$y = f_2(input_{j_3}^3)$$

Backward Pass

$$E = f_3(z, y)$$

$$w_{ij}^{3'} = w_{ij}^3 - \lambda \frac{\partial E}{\partial w^3}$$

$$w_{ij}^{2'} = w_{ij}^2 - \lambda \frac{\partial E}{\partial w^2}$$

$$w_{ij}^{1'} = w_{ij}^1 - \lambda \frac{\partial E}{\partial w^1}$$