Machine Learning
Homework 1
Haoyu Wu

1. Backpropagation

$$a_j^{(\ell)} = g_\ell(z_j^{(\ell)})$$

$$z_j^{(\ell)} = \sum_{i=1}^{n_{\ell-1}} a_i^{(\ell-1)} W_{ij}^{(\ell)} + b_j^{(\ell)}$$

$$C = \frac{1}{|D|}\sum_{x\in D} c(x)$$

$$\delta_j^{(\ell)} = \frac{\partial c}{\partial z_j^{(\ell)}}$$

a) $$\delta_j^{(\ell)} = \frac{\partial c}{\partial z_j^{(\ell)}} = \frac{\partial c}{\partial a_j^{(\ell)}}\frac{\partial a_j^{(\ell)}}{\partial z_j^{(\ell)}} = \frac{\partial c}{\partial a_j^{(\ell)}}\frac{\partial g_\ell(z_j^{(\ell)})}{\partial z_j^{(\ell)}} = \frac{\partial c}{\partial a_j^{(\ell)}} g_\ell'(z_j^{(\ell)})$$

For $\ell = L$, $\delta_j^{(L)} = \frac{\partial c}{\partial a_j^{(L)}} g_L'(z_j^{(L)})$

For $\ell = L-1$, $$\frac{\partial c}{\partial a_j^{(L-1)}} = \frac{\partial z_k^{(L)}}{\partial a_j^{(L-1)}}\frac{\partial c}{\partial z_k^{(L)}} = \sum_{k=1}^{n_L} W_{jk}^{(L)}\delta_k^{(L)} = \sum_{k=1}^{n_L}\delta_k^{(L)} W_{kj}^{(L)T}$$

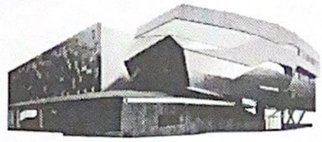
where $$\frac{\partial z_k^{(L)}}{\partial a_j^{(L-1)}} = \frac{\partial}{\partial a_j^{(L-1)}}\left[\sum_{j=1}^{n_{L-1}} a_j^{(L-1)} W_{jk}^{(L)} + b_k^{(L)}\right] = \sum_{j=1}^{n_{L-1}} W_{jk}^{(L)}$$

$$\frac{\partial c}{\partial z_k^{(L)}} = \delta_k^{(L)}$$

So $$\delta_j^{(L-1)} = \frac{\partial c}{\partial a_j^{(L-1)}} g_{L-1}'(z_j^{(L-1)})$$

$$= \sum_{k=1}^{n_L}\delta_k^{(L)} W_{kj}^{(L)T} g_{L-1}'(z_j^{(L-1)})$$

If we back-propagates, ~~for~~ $\ell < L$, $\delta_j^{(\ell)} = \sum_{k=1}^{n_{\ell+1}}\delta_k^{(\ell+1)} W_{kj}^{(\ell+1)T} g_\ell'(z_j^{(\ell)})$



b) $$\frac{\partial C}{\partial W_{ij}^{(\ell)}} = \frac{\partial C}{\partial z_j^{(\ell)}} \frac{\partial z_j^{(\ell)}}{\partial W_{ij}^{(\ell)}}$$

$$= \delta_j^{(\ell)} \frac{\partial}{\partial W_{ij}^{(\ell)}} \left[\sum_{i=1}^{n_{\ell-1}} a_i^{(\ell-1)} W_{ij}^{(\ell)} + b_j^{(\ell)} \right]$$

$$= \delta_j^{(\ell)} a_i^{(\ell-1)}$$

$$= a_i^{(\ell-1)} \delta_j^{(\ell)}$$

$$\frac{\partial C}{\partial b_j^{(\ell)}} = \frac{\partial C}{\partial z_j^{(\ell)}} \frac{\partial z_j^{(\ell)}}{\partial b_j^{(\ell)}}$$

$$= \delta_j^{(\ell)} \frac{\partial}{\partial b_j^{(\ell)}} \left[\sum_{i=1}^{n_{\ell-1}} a_i^{(\ell-1)} W_{ij}^{(\ell)} + b_j^{(\ell)} \right]$$

$$= \delta_j^{(\ell)}$$