PERIMETER PHYSICS

Machine Learning Homework 1 Havyu Wu

If we back-propagator,
$$A_{j}^{(e)} = g_{\ell}(z_{j}^{(e)})$$

$$Z_{j}^{(e)} = g_{\ell}(z_{j}^{(e)})$$

$$Z_{j}^{(e)} = \sum_{l=1}^{n_{\ell-1}} \alpha_{i}^{(\ell-1)} W_{ij}^{(l)} + b_{j}^{(\ell)}$$

$$C = \frac{1}{|D|} \sum_{k \in D} c_{\ell}(x)$$

$$S_{j}^{(e)} = \frac{3c}{3z_{j}^{(e)}} = \frac{3c}{3\alpha_{j}^{(e)}} \frac{3\alpha_{j}^{(e)}}{3z_{j}^{(e)}} = \frac{3c}{3\alpha_{j}^{(e)}} \frac{3g_{\ell}(z_{j}^{(e)})}{3z_{j}^{(e)}} = \frac{3c}{3\alpha_{j}^{(e)}} g_{\ell}^{(e)}(z_{j}^{(e)})$$

For $\ell = \ell$, $S_{j}^{(e)} = \frac{3c}{3\alpha_{j}^{(e)}} g_{\ell}^{(e)}(z_{j}^{(e)})$

For $\ell = \ell$, $S_{j}^{(e)} = \frac{3c}{3\alpha_{j}^{(e)}} g_{\ell}^{(e)}(z_{j}^{(e)})$

$$V_{jk}^{(e)} = \frac{3c}{3\alpha_{j}^{(e)}} \frac{3c}{3\alpha_{j}^{(e)}} \sum_{j=1}^{n_{\ell-1}} W_{jk}^{(e)} G_{k}^{(e)} = \sum_{k=1}^{n_{\ell-1}} G_{k}^{(e)} W_{kj}^{(e)}$$

$$V_{jk}^{(e)} = \frac{3c}{3\alpha_{j}^{(e)}} g_{\ell}^{(e)}(z_{j}^{(e)})$$

$$S_{j}^{(e)} = \frac{3c}{3\alpha_{j}^{(e)}} g_{\ell}^{(e)}(z_{j}^{(e)})$$

$$= \sum_{k=1}^{n_{\ell-1}} G_{k}^{(e)}(z_{j}^{(e)})$$

$$= \sum_{k=1}^{n_{\ell-1}} G_{k}^{(e)}(z_{j}^{(e)})$$

If we back-propagator, for $\ell = \ell < \ell$, $S_{j}^{(e)} = \sum_{k=1}^{n_{\ell-1}} G_{k}^{(e+1)} W_{kj}^{(e+1)} f_{\ell}^{(e)}(z_{j}^{(e)})$





www.perlmeterInstItute.ca

Phone (519) 569-7600 Fax (519) 569-7611

PERIMETER PINSTITUTE FOR THEORETICAL PHYSICS

b)
$$\frac{\partial c}{\partial W_{ij}^{(\ell)}} = \frac{\partial c}{\partial z_{j}^{(\ell)}} \frac{\partial z_{j}^{(\ell)}}{\partial W_{ij}^{(\ell)}}$$

$$= \int_{j}^{(\ell)} \frac{\partial}{\partial W_{ij}^{(\ell)}} \left[\sum_{i=1}^{n_{\ell-1}} a_{i}^{(\ell-1)} W_{ij}^{(\ell)} + b_{j}^{(\ell)} \right]$$

$$= \int_{j}^{(\ell)} a_{i}^{(\ell-1)}$$

$$= a_{i}^{(\ell-1)} f_{j}^{(\ell)}$$

$$\frac{\partial c}{\partial b_{j}^{(\ell)}} = \frac{\partial c}{\partial z_{j}^{(\ell)}} \frac{\partial \overline{z}_{j}^{(\ell)}}{\partial b_{j}^{(\ell)}}$$

$$= \int_{j}^{(\ell)} \frac{\partial}{\partial b_{j}^{(\ell)}} \left[\sum_{i=1}^{n_{\ell-1}} \alpha_{i}^{(\ell-1)} W_{ij}^{(\ell)} + b_{j}^{(\ell)} \right]$$

$$= \int_{j}^{(\ell)} \ell d\ell$$