Backpropagation

November 12, 2016

1 Definitions, Forward Propagation

$$x_{bc} = \text{input} \tag{1}$$
$$t_b = \text{target} \tag{2}$$
$$y_b = \text{output} \tag{3}$$
$$\rho(x) = \text{RELU}(x) = \max(0, x) \tag{4}$$
$$\theta(x) = \text{Heavside}(x) = \rho'(x) \tag{5}$$
$$\sigma_j = \sigma_j(\{x_i\}) = \text{softmax}(\{x_i\}) = e^{x_j} / \sum_i e^{x_i} \tag{6}$$
$$\partial_{x_j}\sigma_i = \sigma_i(\delta_{ij} - \sigma_j) \tag{7}$$
$$m_{fg}(x_{de}) = \text{maxpool}_{fg}(x_{de}) \tag{8}$$
$$m'(x_{de}) = 1 \text{ if } x_{de} \text{ max in stride else } 0 \tag{9}$$
$$z^{(1)}_{de} = \sum_{b,c} [W^{(conv)}_{d-b,e-c} x_{bc}] + b^{(1)}_{de} \tag{10}$$
$$a^{(1)}_{de} = \rho(z^{(1)}_{de}) \tag{11}$$
$$z^{(2)}_{fg} = m_{fg}(a^{(1)}_{de}) + b^{(2)}_{fg} \tag{12}$$
$$a^{(2)}_{fg} = \rho(z^{(2)}_{fg}) \tag{13}$$
$$z^{(3)}_{i} = \sum_{fg} [W^{(full)}_{ifg} a^{(2)}_{fg}] + b^{(3)}_{i} \tag{14}$$
$$y_h = \sigma_h(z^{(3)}_i) = \sigma_h \tag{15}$$
$$\text{label}_h = \text{argmax}(y_h) \tag{16}$$
$$S = -\sum_h t_h ln(y_h) \tag{17}$$
$$\tag{18}$$

2 Simple Derivatives

$$\partial_{y_h} S = -t_h / y_h \tag{19}$$
$$\partial_{z_i^{(3)}} y_h = y_h(\delta_{ih} - y_i) \tag{20}$$
$$\partial_{a_{fg}^{(2)}} z_i^{(3)} = W_{ifg}^{(full)} \tag{21}$$
$$\partial_{z_{fg}^{(2)}} a_{fg}^{(2)} = \theta(z_{fg}^{(2)}) \tag{22}$$
$$\partial_{a_{de}^{(1)}} z_{fg}^{(2)} = m'(a_{de}^{(1)}) \tag{23}$$
$$\partial_{z_{de}^{(1)}} a_{de}^{(1)} = \theta(z_{de}^{(1)}) \tag{24}$$
$$\partial_{x_{bc}} z_{de}^{(1)} = W_{d-b,e-c}^{(conv)} \tag{25}$$
$$\partial_{W_{afg}^{(full)}} z_i^{(3)} = \delta_{ia} a_{fg}^{(2)} \tag{26}$$
$$\partial_{W_{w-x,y-z}^{(conv)}} z_{de}^{(1)} = x_{d-w+x,e-y+z} \tag{27}$$
$$\partial_{W_{yz}^{(conv)}} z_{de}^{(1)} = x_{d-y,e-z} \tag{28}$$

3 Parameter Derivatives

$$\frac{\partial S}{\partial W_{afg}^{(full)}} \tag{29}$$
$$= \sum_h \frac{\partial S}{\partial y_h} \sum_i \frac{\partial y_h}{\partial z_i^{(3)}} \frac{\partial z_i^{(3)}}{\partial W_{afg}^{(full)}} \tag{30}$$
$$= -\sum_h \frac{t_h}{y_h} \sum_i y_h(\delta_{ih} - \sigma_i)\delta_{ia} a_{fg}^{(2)} \tag{31}$$
$$= -\sum_h t_h(\delta_{ah} - y_a) a_{fg}^{(2)} \tag{32}$$
$$= a_{fg}^{(2)} \sum_h t_h(y_a - \delta_{ah}) \tag{33}$$
$$= a_{fg}^{(2)} \left[y_a \sum_h t_h - \sum_h t_h \delta_{ah}) \right] \tag{34}$$
$$= a_{fg}^{(2)}(y_a - t_a) \tag{35}$$

program@epstopd

$$\frac{\partial S}{\partial W_{yz}^{(conv)}} \tag{36}$$
$$= \sum_h \frac{\partial S}{\partial y_h} \sum_i \frac{\partial y_h}{\partial a_i^{(3)}} \frac{\partial a_i^{(3)}}{\partial z_i^{(3)}} \sum_{fg} \frac{\partial z_i^{(3)}}{\partial a_{fg}^{(2)}} \frac{\partial a_{fg}^{(2)}}{\partial z_{fg}^{(2)}} \sum_{de} \frac{\partial z_{fg}^{(2)}}{\partial a_{de}^{(1)}} \frac{\partial a_{de}^{(1)}}{\partial z_{de}^{(1)}} \sum_{bc} \frac{\partial [W_{d-b,e-c}^{(conv)} x_{bc}]}{\partial W_{yz}^{(conv)}} \tag{37}$$

$$= -\sum_h \frac{t_h}{y_h} \sum_i \sigma_h(\delta_{ih} - \sigma_i) \sum_{fg} \frac{\partial z_i^{(3)}}{\partial a_{fg}^{(2)}} \theta(z_{fg}^{(2)}) \sum_{de} \frac{\partial z_{fg}^{(2)}}{\partial a_{de}^{(1)}} \theta(z_{de}^{(1)}) x_{d-y,e-z} \tag{38}$$

$$= -\sum_h \frac{t_h}{y_h} \sum_i \sigma_h(\delta_{ih} - \sigma_i) \sum_{fg} W_{ifg}^{(full)} \theta(z_{fg}^{(2)}) \sum_{de} m'_{fg}(a_{de}^{(1)}) \theta(z_{de}^{(1)}) x_{d-y,e-z} \tag{39}$$

$$= -\sum_h t_h \sum_i (\delta_{ih} - y_i) \sum_{fg} W_{ifg}^{(full)} \theta(z_{fg}^{(2)}) \sum_{de} m'_{fg}(a_{de}^{(1)}) \theta(z_{de}^{(1)}) x_{d-y,e-z} \tag{40}$$

$$\frac{\partial S}{\partial b_j^{(3)}} \tag{41}$$

$$= \sum_h \frac{\partial S}{\partial y_h} \sum_i \frac{\partial y_h}{\partial a_i^{(3)}} \frac{\partial a_i^{(3)}}{\partial z_i^{(3)}} \frac{\partial z_i^{(3)}}{\partial b_j^{(3)}} \tag{42}$$

$$= -\sum_h \frac{t_h}{y_h} \sum_i \sigma_h(\delta_{ih} - \sigma_i) \delta_{ij} \tag{43}$$

$$= -\sum_h t_h(\delta_{jh} - y_j) \tag{44}$$

$$= y_j - t_j \tag{45}$$

$$\frac{\partial S}{\partial b_{lm}^{(2)}} \tag{46}$$

$$= \sum_h \frac{\partial S}{\partial y_h} \sum_i \frac{\partial y_h}{\partial z_i^{(3)}} \sum_{fg} \frac{\partial z_i^{(3)}}{\partial a_{fg}^{(2)}} \frac{\partial a_{fg}^{(2)}}{\partial z_{fg}^{(2)}} \delta_{fl} \delta_{gm} \tag{47}$$

$$= -\sum_h \frac{t_h}{y_h} \sum_i \sigma_h(\delta_{ih} - \sigma_i) \sum_{fg} W_{ifg}^{(full)} \theta(z_{fg}^{(2)}) \delta_{fl} \delta_{gm} \tag{48}$$

$$= -\sum_h t_h \sum_i y_h(\delta_{ih} - y_i) W_{ilm}^{(full)} \theta(z_{lm}^{(2)}) \tag{49}$$

$$\frac{\partial S}{\partial b_{lm}^{(1)}} \tag{50}$$

$$= \sum_h \frac{\partial S}{\partial y_h} \sum_i \frac{\partial y_h}{\partial a_i^{(3)}} \frac{\partial a_i^{(3)}}{\partial z_i^{(3)}} \sum_{fg} \frac{\partial z_i^{(3)}}{\partial a_{fg}^{(2)}} \frac{\partial a_{fg}^{(2)}}{\partial z_{fg}^{(2)}} \sum_{de} \frac{\partial z_{fg}^{(2)}}{\partial a_{de}^{(1)}} \frac{\partial a_{de}^{(1)}}{\partial z_{de}^{(1)}} \delta_{dl} \delta_{em} \tag{51}$$

$$= -\sum_h \frac{t_h}{y_h} \sum_i \sigma_h(\delta_{ih} - \sigma_i) \sum_{fg} W_{ifg}^{(full)} \theta(z_{fg}^{(2)}) \sum_{de} m'_{fg}(a_{de}^{(1)}) \theta(z_{de}^{(1)}) \delta_{dl} \delta_{em} \tag{52}$$

$$= -\sum_h \frac{t_h}{y_h} \sum_i \sigma_h(\delta_{ih} - \sigma_i) \sum_{fg} W_{ifg}^{(full)} \theta(z_{fg}^{(2)}) m'_{fg}(a_{lm}^{(1)}) \theta(z_{lm}^{(1)}) \tag{53}$$

$$= -\sum_h t_h \sum_i (\delta_{ih} - y_i) \sum_{fg} W_{ifg}^{(full)} \theta(z_{fg}^{(2)}) m'_{fg}(a_{lm}^{(1)}) \theta(z_{lm}^{(1)}) \tag{54}$$

$$\tag{55}$$