CS542 ML PSet3

Tuesday, March 3, 2020 9:08 PM

Problem 1

$f(z) = 1/1+\exp(-z)$

$L(\hat{y},y) = -y\ln(\hat{y}) - (1-y)\ln(1-\hat{y})$

$h^{(0)} = x \rightarrow a^{(1)} = b^{(1)} + W^{(1)}h^{(0)}$

$\downarrow$

$h^{(1)} = f(a^{(1)}) = 1/1+\exp(-a^{(1)})$

$\downarrow$

$a^{(2)} = b^{(2)} + W^{(2)}h^{(1)} \rightarrow h^{(2)} = 1/1+\exp(-a^{(2)})$

$\hookrightarrow \hat{y} = h^{(2)}$

$J = L(\hat{y},y) + \lambda\Omega(\theta)$

$$\nabla_{\hat{y}} J = \nabla_{\hat{y}} L(\hat{y},y) = \frac{\delta}{\delta\hat{y}}\left[-y\ln(\hat{y}) - (1-y)\ln(1-\hat{y})\right]$$

$$= \frac{-y}{\hat{y}} - \frac{(1-y)}{(1-\hat{y})}(-1) = \frac{(1-y)}{(1-\hat{y})} - \frac{y}{\hat{y}}$$

$$= \frac{\hat{y}(1-y) - y(1-\hat{y})}{\hat{y}(1-\hat{y})} = \frac{\hat{y} - \cancel{\hat{y}y} - y + \cancel{\hat{y}y}}{\hat{y}(1-\hat{y})}$$

$$= \frac{\hat{y}-y}{\hat{y}(1-\hat{y})} = \frac{h^{(2)}-y}{(1-h^{(2)})h^{(2)}} \Rightarrow g \leftarrow \nabla_{\hat{y}} J$$

$$\nabla_{a^{(2)}} J = \nabla_{a^{(2)}} L(\hat{y},y) = \left(\frac{\delta}{\delta\hat{y}} L(\hat{y},y)\right)\cdot\left(\frac{\delta}{\delta a^{(2)}}\hat{y}\right)$$

$$\frac{\delta}{\delta a^{(2)}}\hat{y} = \frac{\delta}{\delta a^{(2)}} h^{(2)} = f'(a^{(2)})$$

$$= \frac{-1}{(1+\exp(-a^{(2)}))^2}\cdot\frac{\delta}{\delta a^{(2)}}(1+\exp(-a^{(2)}))$$

$$= \frac{\exp(-a^{(2)})}{(1+\exp(-a^{(2)}))^2} = (h^{(2)})^2(\exp(-a^{(2)}))$$

$$= (h^{(2)})^2\left(\frac{1}{h^{(2)}} - 1\right) = h^{(2)} - (h^{(2)})^2$$

$$= h^{(2)}(1-h^{(2)})$$

$$* f'(z) = \cancel{-}f(z)^2\cdot(\exp(-z))\cdot\cancel{(-1)}$$
$$= f(z)^2\cdot\left(\frac{1}{f(z)} - 1\right)$$
$$= f(z)(1-f(z))$$

$$\hookrightarrow \nabla_{a^{(2)}} J = \frac{h^{(2)}-y}{\cancel{(1-h^{(2)})h^{(2)}}}\cdot\cancel{h^{(2)}(1-h^{(2)})} = h^{(2)} - y$$

$g \leftarrow \nabla_{a^{(2)}} J$

$$\nabla_{b^{(2)}} J = g + \cancel{\lambda\nabla_{b^{(2)}}\Omega(\theta)}^{\,0} = g = h^{(2)} - y$$

$$\nabla_{W^{(2)}} J = g h^{(1)T} + \cancel{\lambda\nabla_{W^{(2)}}\Omega(\theta)}^{\,0} = (h^{(2)}-y)h^{(1)T}$$

$$\nabla_{h^{(1)}} J = W^{(2)T} g = W^{(2)T}(h^{(2)}-y) \rightarrow \text{scalar}$$

$$= (h^{(2)}-y)W^{(2)T}$$

$g \leftarrow \nabla_{h^{(1)}} J$

$$\nabla_{a^{(1)}} J = g \odot f'(a^{(1)}) = g \odot h^{(1)} \odot (1-h^{(1)})$$

* element-wise for vector case

$$= (h^{(2)}-y)W^{(2)T} \odot (h^{(1)} - \cancel{h^{(1)}\odot h^{(1)}}^{\,[h^{(1)}]^2})$$

$$\nabla_{b^{(1)}} J = \nabla_{a^{(1)}} J$$

$$\nabla_{W^{(1)}} J = \nabla_{a^{(1)}} J\, h^{(0)T} = (h^{(2)}-y)\left[W^{(2)T}\odot(h^{(1)} - [h^{(1)}]^2)\right]h^{(0)T}$$