Live Lecture Feedback App

•••

Derek Law
Felix Chiu
Jiranart Vacheesuthum
Vasin Wongrassamee

Hidden Layer Cost Gradients 2

• Similarly, for hidden layer L-2 (if there is one) we get

$$\frac{\partial C}{\partial w_{jk}^{L-2}} = \frac{\partial z_{j}^{L-2}}{\partial w_{jk}^{L-2}} \cdot \frac{\partial a_{j}^{L-2}}{\partial z_{ij}^{L-2}} \cdot \sum_{i1} \left(\frac{\partial z_{i1}^{L-1}}{\partial a_{j}^{L-2}} \cdot \frac{\partial a_{i1}^{L-1}}{\partial z_{i1}^{L-1}} \cdot \sum_{i2} \left(\frac{\partial z_{i2}^{L}}{\partial a_{i1}^{L-1}} \cdot \frac{\partial a_{i2}^{L}}{\partial z_{i2}^{L}} \cdot \frac{\partial C}{\partial a_{i2}^{L}} \right) \right)$$

Differentiating, we get

$$\frac{\partial C}{\partial w_{jk}^{L-2}} = a_k^{L-3} \sigma'(z_j^{L-2}) \sum_{i1} \left(w_{i1j}^{L-1} \sigma'(z_{i1}^{L-1}) \sum_{i2} \left(w_{i2i1}^{L} \sigma'(z_{i2}^{L}) \frac{\partial C}{\partial a_{i2}^{L}} \right) \right)$$







