

# Softmax loss

$$L_i = -\log\left(\frac{e^{f_{y_i}}}{\sum_j e^{f_j}}\right) = -f_{y_i} + \log\left(\sum_j e^{f_j}\right)$$

$$f_{y_i} = x_i w_j$$

$$\nabla_{w_j} L_i = -x_i + \frac{e^{f_{y_i}}}{\sum_j e^{f_j}} x_i$$

correct class.

$$\nabla_{w_j} L_i = \frac{e^{f_j} x_i}{\sum_j e^{f_j}}$$

Rest