6.867 Problem Set 3

November 4, 2015

Neural Networks

$$\frac{\partial J(w)}{\partial w_{kj}^{(2)}} = \sum_{i=1}^{N} z_j \left(y_k^{(i)}(1 - \sigma(a_k^{(2)})) + (1 - y_k^{(i)})(\sigma(a_k^{(2)})) \right) + 2\lambda w_{kj}^{(2)} \tag{1}$$

$$\frac{\partial J(w)}{\partial w_{jd}^{(1)}} = \sum_{i=1}^{N} x_i \sum_{k=1}^{K} \left(y_k^{(i)}(1 - \sigma(a_k^{(2)})) + (1 - y_k^{(i)})(\sigma(a_k^{(2)})) \right) w_{kj}^{(2)} \left(\frac{1}{1 + e^{-a_{jd}^{(1)}}} \right) \left(1 - \frac{1}{1 + e^{-a_{jd}^{(1)}}} \right) + 2\lambda w_{jd}^{(1)} \tag{2}$$