Question 2: continued

(input layer derivative)

$$\frac{\partial R}{\partial w_{ij}} = \frac{1}{N}\sum_{n}\left[\frac{\partial L^n}{\partial a_j^n}\right]\left(\frac{\partial a_j^n}{\partial w_{ij}}\right) = \frac{1}{N}\sum_{n}\left[\sum_{\ell}\frac{\partial L^n}{\partial a_k^n}\frac{\partial a_k^n}{\partial a_j^n}\right]\left(\frac{\partial a_j^n}{\partial w_{ij}}\right) = \frac{\sum_{n}\delta_j^n z_i^n}{N}$$

where δ_j^n defined from $= \frac{1}{N}\sum_{n}\left[\sum_{\ell}\delta_k^n \frac{\partial a_k^n}{\partial a_j^n}\right] z_i^n = \frac{1}{N}\sum_{n}\underbrace{\left[\sum_{\ell}\delta_k^n w_{ij} f'(a_j^n;\theta)\right]}_{\delta_j^n} z_i^n$

where z_i^n just our initial inputs x_1 x_2

Steps then: [Gradient Descent]

$$w_{ij}^{t+1} = w_{ij}^{t} - \eta\frac{\partial R}{\partial w_{ij}} \qquad w_{jk}^{t+1} = w_{jk}^{t} - \eta\frac{\partial R}{\partial w_{jk}} \qquad w_{k\ell}^{t+1} = w_{k\ell}^{t} - \eta\frac{\partial R}{\partial w_{k\ell}}$$

with derivatives as defined above.