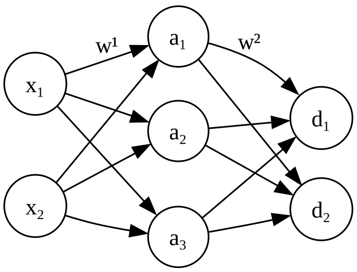
Observations

$$i, j, k, l, L, m, M, n, N, o \in \quad \mathcal{N} \tag{1}$$
$$X \in \quad \mathcal{R}^{n\times o} \tag{2}$$
$$Y \in \quad \mathcal{R}^{n\times m} \tag{3}$$

Neural Network



$$a^0 = \quad x_{1\times p}(n) \tag{4}$$
$$a^L = \quad d_{1\times m}(n) \tag{5}$$
$$a^l = \quad \varphi(z^l) \tag{6}$$
$$z^l = \quad a^{l-1}W^l \tag{7}$$

Gradient Descent

$$e(n) = \quad y(n) - d(n) \tag{8}$$
$$\xi(n) = \quad \tfrac{1}{2}ee^\top \tag{9}$$
$$\xi(n) = \quad \tfrac{1}{2}\textstyle\sum_{j=1}^{M}(e_j(n))^2 \tag{10}$$
$$W_{(k+1)} = \quad W_{(k)} - \nabla_W \xi(d, y) \tag{11}$$
$$\xi_{avg}(n) = \quad \tfrac{1}{2n}\textstyle\sum_{n=1}^{N}\sum_{j=1}^{M}(e_j(n))^2 \tag{12}$$
$$\tag{13}$$

Backpropagation

$$\frac{\partial \xi}{\partial \omega_{ij}^l} = \quad \delta_j^l \frac{\partial z_j^l}{\partial \omega_{ij}} \tag{14}$$
$$\delta_j^l = \quad \frac{\partial \xi}{\partial z_j^l} \tag{15}$$
$$\frac{\partial z_j^l}{\partial \omega_{ij}} = \quad a_i^{l-1} \tag{16}$$
$$\tag{17}$$

Output Layer

$$\delta_j^L = \quad \frac{\partial \xi}{\partial z_j^L} = \frac{\partial \xi}{\partial a_j^L}\frac{\partial a_j^L}{\partial z_j^L} \tag{18}$$
$$\delta_j^L = \quad \frac{\partial \xi}{\partial a_j^L}\dot{\varphi}(z_j^L) \tag{19}$$
$$= \quad -e_j\dot{\varphi}(z_j^L) \tag{20}$$

Hidden Layer

$$\delta_j^l = \quad \frac{\partial \xi}{\partial z_j^l} = \sum_k \frac{\partial \xi}{\partial z_k^{l+1}}\frac{\partial z_k^{l+1}}{\partial z_j^l} \tag{21}$$
$$\delta_j^l = \quad \sum_k \delta_k^{l+1}\frac{\partial z_k^{l+1}}{\partial z_j^l} \tag{22}$$
$$\frac{\partial z_k^{l+1}}{\partial z_j^l} = \quad \frac{\partial}{\partial z_j^l}\left(\sum_j \omega_{jk}^{l+1}\varphi(z_j^l)\right) \tag{23}$$
$$\frac{\partial z_k^{l+1}}{\partial z_j^l} = \quad \omega_{jk}\dot{\varphi}(z_j^l) \tag{24}$$
$$\delta_j^l = \quad \sum_k \delta_k^{l+1}\omega_{jk}^{l+1}\dot{\varphi}(z_j^l) \tag{25}$$
$$\tag{26}$$