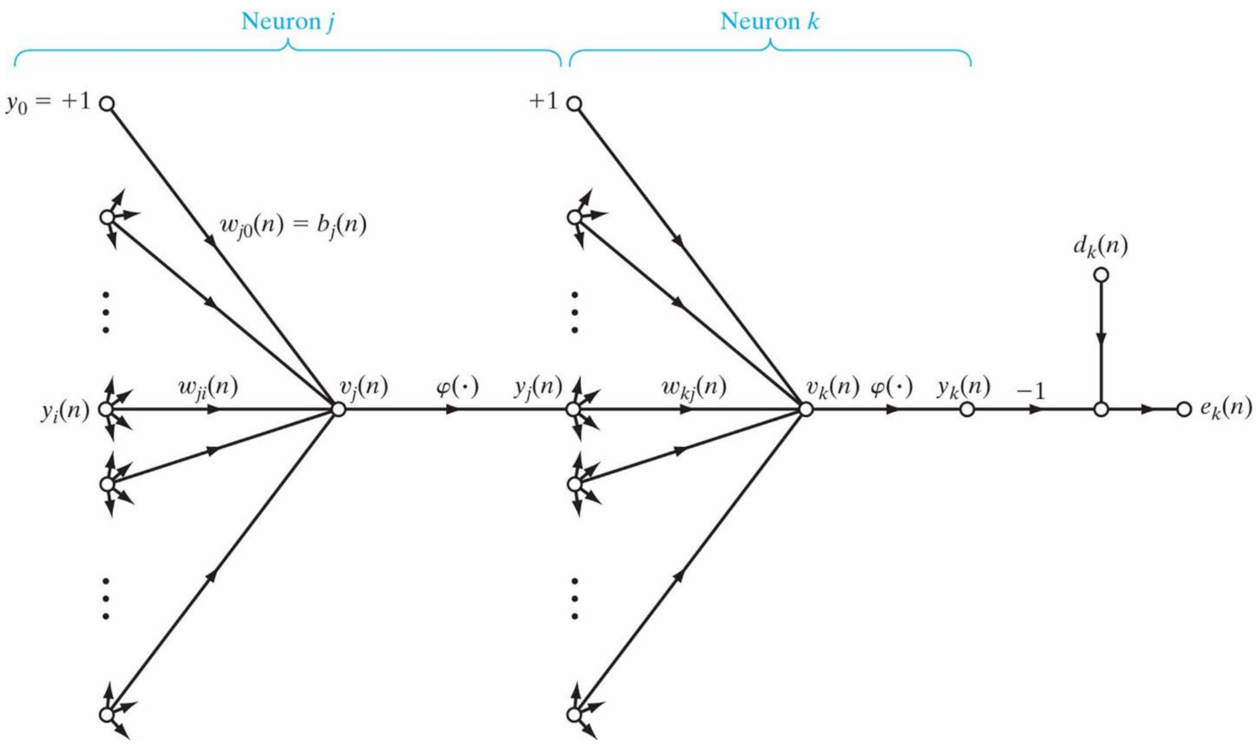
Consider the following signal-flow graph of a fully-connected neural network that consists of an input layer, one hidden layer and an output layer. y_i is the i^{th} input node in the input layer. Neuron j is the j^{th} neuron in the hidden layer and neuron k is the k^{th} output neuron. Assume the activation function $\varphi(.)$ is sigmoid.



(a). Using back-propagation on the output neuron k show that the weight correction Δw_{kj} for the n^{th} iteration is given by

$$\Delta w_{kj}(n) = \eta . \delta_k(n) . y_j(n)$$

Where η is the learning rate and the local gradient $\delta_k(n) = [d_k(n) - y_k(n)] . \left[y_k(n)\left(1 - y_k(n)\right)\right]$

(b). Using back-propagation on the hidden neuron j show that the weight correction Δw_{ji} for the n^{th} iteration is given by

$$\Delta w_{ji}(n) = \eta . \delta_j(n) . y_i(n)$$

Where the $\delta_j(n)$ is the overall backpropagated gradient from the layer to the immediate right (i.e., the output layer) given by

$$\delta_j(n) = \sum_k \delta_k(n) . w_{kj}(n) . y_j(n) . \left(1 - y_j(n)\right)$$

and $y_j(n)$ is the output of the hidden neuron j.

Notice that the effect of all $e_k{}'$s must be included, hence the summation over k.