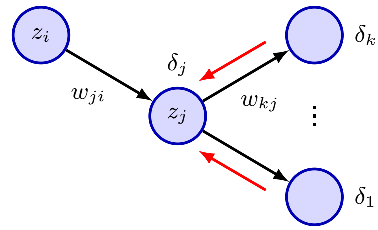

Figure 8.1 Illustration of the calculation of δ_j for hidden unit j by backpropagation of the δ's from those units k to which unit j sends connections. The black arrows denote the direction of information flow during forward propagation, and the red arrows indicate the backward propagation of error information.

where the δ's are often referred to as *errors* for reasons we will see shortly. Using (8.5), we can write

$$\frac{\partial a_j}{\partial w_{ji}} = z_i. \tag{8.9}$$

Substituting (8.8) and (8.9) into (8.7), we then obtain

$$\frac{\partial E_n}{\partial w_{ji}} = \delta_j z_i. \tag{8.10}$$

Equation (8.10) tells us that the required derivative is obtained simply by multiplying the value of δ for the unit at the output end of the weight by the value of z for the unit at the input end of the weight (where $z = 1$ for a bias). Note that this takes the same form as that found for the simple linear model in (8.4). Thus, to evaluate the derivatives, we need calculate only the value of δ_j for each hidden and output unit in the network and then apply (8.10).

As we have seen already, for the output units, we have

$$\delta_k = y_k - t_k \tag{8.11}$$

Section 5.4.6

provided we are using the canonical link as the output-unit activation function. To evaluate the δ's for hidden units, we again make use of the chain rule for partial derivatives:

$$\delta_j \equiv \frac{\partial E_n}{\partial a_j} = \sum_k \frac{\partial E_n}{\partial a_k}\frac{\partial a_k}{\partial a_j} \tag{8.12}$$

where the sum runs over all units k to which unit j sends connections. The arrangement of units and weights is illustrated in Figure 8.1. Note that the units labelled k include other hidden units and/or output units. In writing down (8.12), we are making use of the fact that variations in a_j give rise to variations in the error function only through variations in the variables a_k.

If we now substitute the definition of δ_j given by (8.8) into (8.12) and make use of (8.5) and (8.6), we obtain the following *backpropagation* formula:

Exercise 8.1

$$\delta_j = h'(a_j)\sum_k w_{kj}\delta_k, \tag{8.13}$$

which tells us that the value of δ for a particular hidden unit can be obtained by propagating the δ's backwards from units higher up in the network, as illustrated