labeled training sample (x, y)

$$x \coloneqq \begin{pmatrix} i_1 \\ i_2 \end{pmatrix} = \begin{pmatrix} 0.05 \\ 0.10 \end{pmatrix} \qquad y \coloneqq \begin{pmatrix} o_1 \\ o_2 \end{pmatrix} = \begin{pmatrix} 0.01 \\ 0.99 \end{pmatrix}$$

initial weight matrices W_1 and W_2

$$W_1^{(0)} \coloneqq \begin{pmatrix} w_1 & w_2 \\ w_3 & w_4 \end{pmatrix} = \begin{pmatrix} 0.15 & 0.20 \\ 0.25 & 0.30 \end{pmatrix} \qquad W_2^{(0)} \coloneqq \begin{pmatrix} w_5 & w_6 \\ w_7 & w_8 \end{pmatrix} = \begin{pmatrix} 0.40 & 0.45 \\ 0.50 & 0.55 \end{pmatrix}$$

initial bias vectors B_1 and B_2

$$B_1 \coloneqq \begin{pmatrix} b_1 \\ b_1 \end{pmatrix} = \begin{pmatrix} 0.35 \\ 0.35 \end{pmatrix} \qquad B_2 \coloneqq \begin{pmatrix} b_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} 0.60 \\ 0.60 \end{pmatrix}$$

net input to first layer t_1

$$t_1 = W_1 x + B_1 = \begin{pmatrix} 0.3775 \\ 0.3925 \end{pmatrix}$$

output of first layer f_1

$$f_1 = \sigma(t_1) = \begin{pmatrix} 0.593270 \\ 0.596884 \end{pmatrix}$$

net input to second layer t_2

$$t_2 = W_2 f_1 + B_2 = \begin{pmatrix} 1.10591 \\ 1.22492 \end{pmatrix}$$

output of second layer f_2

$$f_2 = \sigma(t_2) = \begin{pmatrix} 0.751365 \\ 0.772928 \end{pmatrix}$$

total error e_0

$$e_0 = \frac{1}{2} \left\| y - f_2 \right\|^2 = 0.298371$$

weight matrices after one backpropagation step $W_1^{(1)}, W_2^{(1)}$ without changing bias

$$W_1^{(0)} = \begin{pmatrix} 0.149781 & 0.199561 \\ 0.249751 & 0.299502 \end{pmatrix} \qquad W_2^{(0)} = \begin{pmatrix} 0.358916 & 0.408666 \\ 0.511301 & 0.56137 \end{pmatrix}$$

total error after one backpropagation step and a second forward feed:

$$e_1 = 0.291028$$

$$\implies e_1 < e_0$$