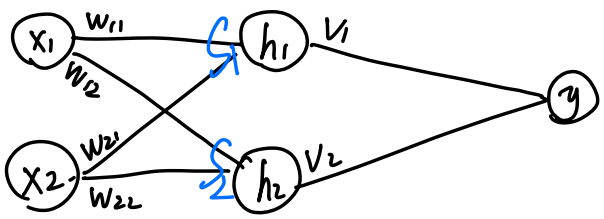
23- S1 - Q4

Q(a) MLP diagram → ref topic 10 P71

input hidden output



(b) ① the net input to each hidden neuron

$$S_1 = w_{11} x_1 + w_{21} x_2 + b_1$$

$$S_2 = w_{12} x_1 + w_{22} x_2 + b_2$$

② σ activation function

$$h_1 = \sigma(S_1) = \frac{1}{1+e^{-S_1}}$$

$$h_2 = \sigma(S_2) = \frac{1}{1+e^{-S_2}}$$

③ output

$$y = V_1 h_1 + V_2 h_2$$

$$= \frac{V_1}{1+e^{-(w_{11} x_1 + w_{21} x_2 + b_1)}} + \frac{V_2}{1+e^{-(w_{12} x_1 + w_{22} x_2 + b_2)}}$$

(c) Q design learning rule → adjust $V=[v_1\ v_2]^T$

Solution

① define the error function as Mean Square Error

$$E=\frac{1}{2}(t-y)^2$$

② compute gradient

$$y=v_1h_1+v_2h_2$$

$$\frac{\partial y}{\partial v_j}=h_j$$

$$\frac{\partial E}{\partial v_j}=-(t-y)\frac{\partial y}{\partial v_j}$$

$$=-(t-y)h_j$$

③ Gradient Descent, η is the learning rate

$$v_j^{new}=v_j^{old}-\eta\frac{\partial E}{\partial v_j}$$

$$=v_j^{old}+\eta(t-y)h_j$$

denote $e=t-y$

$$\Delta v_j=\eta e h_j$$

So $v_j^{new} = v_j^{old} + \delta v_j$