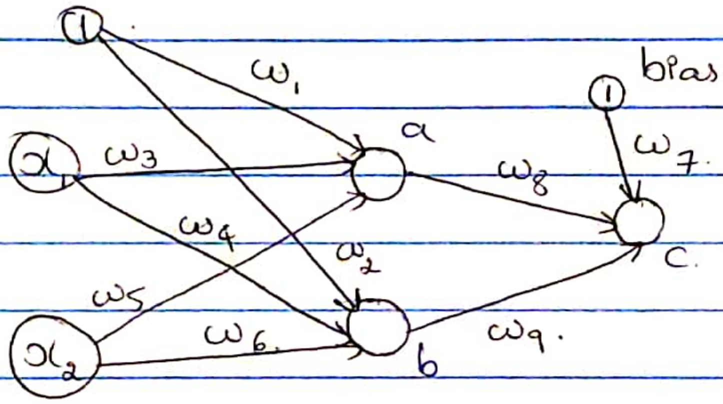
DS5220
Homework-04 : KAVANA VENKATESH

1). Given $a(z) = cz$ and the sigmoid activation function,

$$a(z) = \frac{1}{1+e^{-z}}$$

1.1). The o/p $P(y=1/x, \omega)$ from the given neural n/w. is, calculated as below.



→ o/p @ node a after the application of the linear activation function is,

$$c(\omega_1 + \omega_3 x_1 + \omega_5 x_2).$$

→ o/p at node b will be,

$$c(\omega_2 + \omega_4 x_1 + \omega_6 x_2).$$

→ Finally the o/p @ the node c, which is the output node will be,

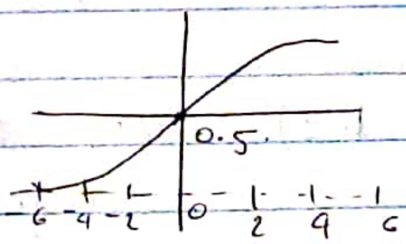
$$\text{Sigmoid}\left(\begin{aligned} &\omega_7 + \omega_8\, C(\omega_1 + \omega_3 x_1 + \omega_5 x_2) \\ &+ \omega_9\, C(\omega_2 + \omega_4 x_1 + \omega_6 x_2) \end{aligned} \right)$$

Hence, $P(y=1/x,\omega) =$

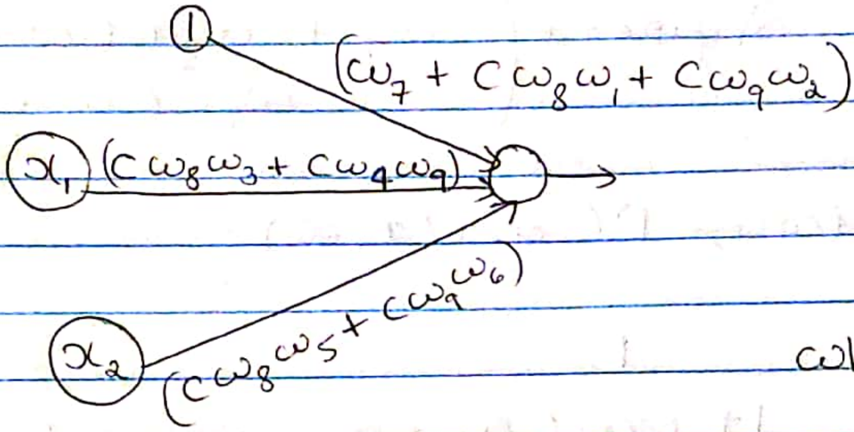
$$\frac{1}{1+\exp\left(\omega_7 + \omega_8 C(\omega_1 + \omega_3 x_1 + \omega_5 x_2) + \omega_9 C(\omega_2 + \omega_4 x_1 + \omega_6 x_2)\right)}$$

→ The classification boundary is declared by the sigmoid function.

Hence, if the final o/p value > 0.5, y = 1
else y = 0.



1.2) The neural net with no hidden layers that is equivalent to the given neural net is,



where,

$$\hat{w} = \begin{bmatrix} c(w_1 w_8 + w_9 w_2) + w_7 \\ w_3 w_8 + w_4 w_9 \\ w_5 w_8 + w_6 w_9 \end{bmatrix}$$ and

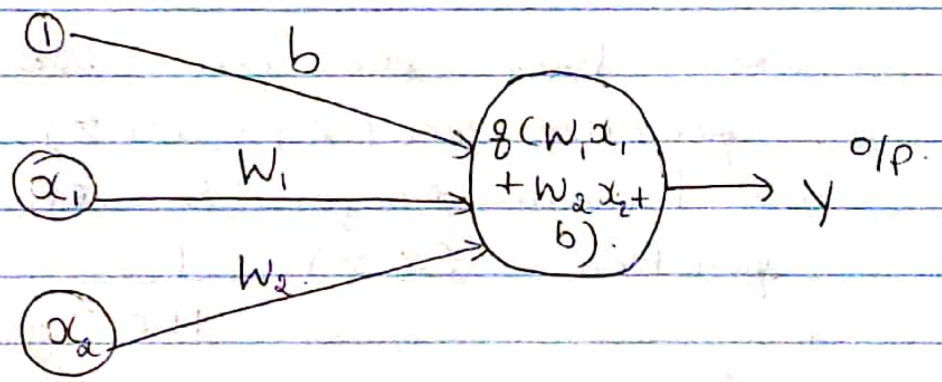
$$\hat{x} = (1 \;\; x_1 \;\; x_2)^T.$$

1.3). Yes. It's true that any multi-layered neural network with only linear activations can be expressed as a neural network without any hidden layer.

This is because any no. of summation and multiplication operations with linear functions will

- always result in a linear function. Hence, the functions in the hidden layers can be replaced by combining those into linear functions at the input layer.

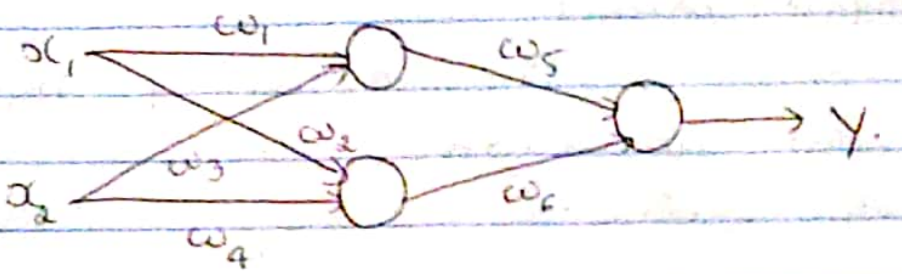
The resulting neural network with no hidden layers will look like this.



Here, W_1 represents the combination of weights for the i/p feature x_1 that come from different hidden layers and W_2 is the combination of weights for x_2.

Scanned with CamScanner

2) To produce the output @ the o/p node in the given form $\frac{1}{1+\exp(\beta_1 x_1+\beta_2 x_2)}$ for the neural n/w.



We have to use linear activation function $a(z) = c(z)$ in the hidden layer & the sigmoid function $S(z) = \frac{1}{1+e^{-z}}$ in the o/p layer.

→ The o/p @ the hidden layer after the application of linear activation function will be,

$$(x_1 \omega_1 + x_2 \omega_3)c + (x_2 \omega_4 + x_1 \omega_2)c = d \text{ (say)}$$

→ At the o/p node,

$$Y = \text{Sigmoid}\left(c(x_1 \omega_1 + x_2 \omega_3) * \omega_5 + c\omega_6 (x_2 \omega_4 + x_1 \omega_2) \right)$$

Reducing the parameter inside the Sigmoid function.

$x_1 c\omega_1\omega_5 + x_2 c\omega_3\omega_5 + x_2 c\omega_4\omega_6 + x_1 c\,\omega_2\omega_6.$

$= (c\omega_1\omega_5 + c\omega_2\omega_6)x_1 + (c\omega_3\omega_5 + c\omega_4\omega_6)x_2.$

$\Rightarrow \beta_1 = c\omega_1\omega_5 + c\,\omega_2\omega_6$ and

$\beta_2 = c\omega_3\omega_5 + c\omega_4\omega_6.$

$$\Rightarrow \boxed{Y = f(X_1, X_2) = \frac{1}{1 + \exp(-(\beta_1 x_1 + \beta_2 x_2))}}$$

or if we consider the values of β_1 and β_2 from within the Sigmoid expression, we can express β_1 and β_2 as,

$$\boxed{\begin{aligned} \beta_1 &= -c(\omega_1\omega_5 + \omega_2\omega_6) \quad \text{and} \\ \beta_2 &= -c(\omega_3\omega_5 + \omega_4\omega_6). \end{aligned}}$$