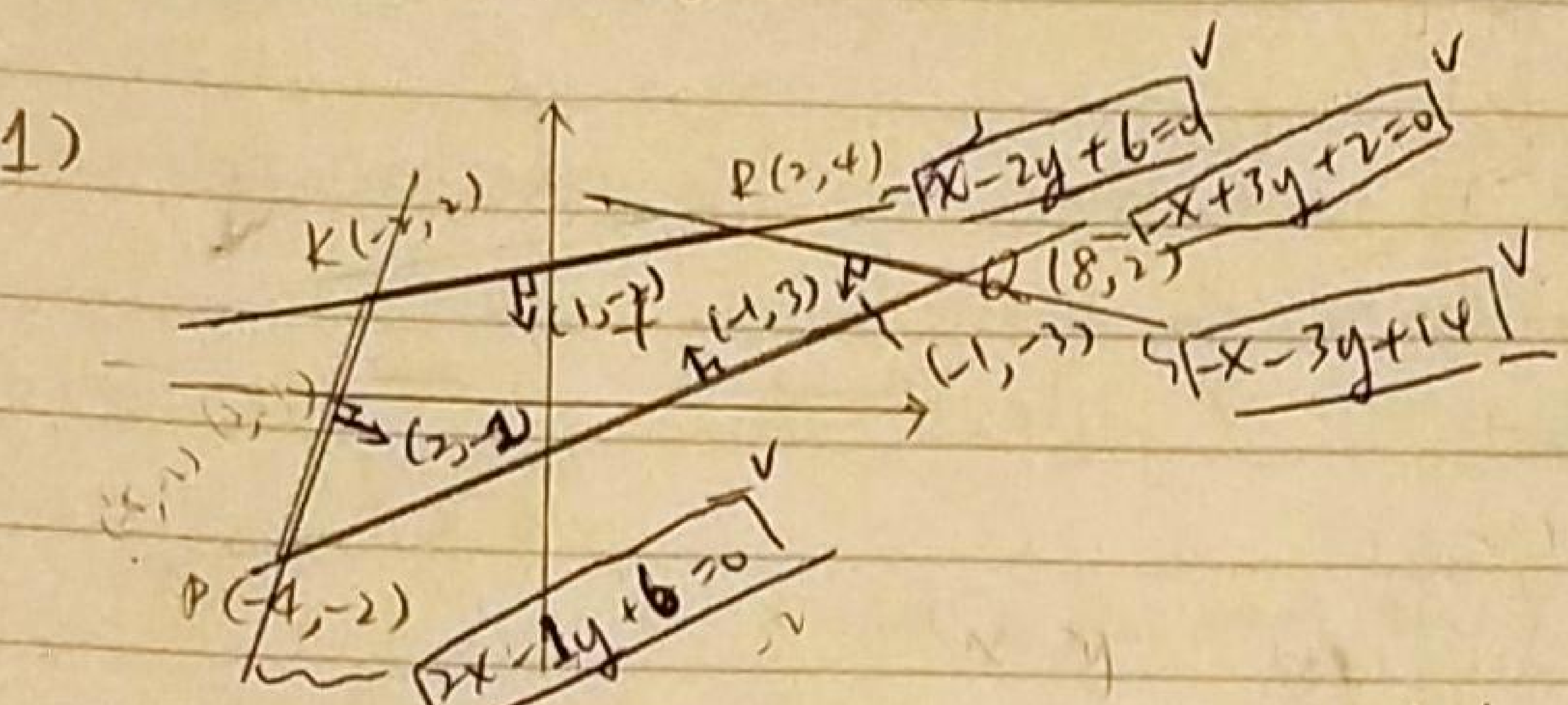
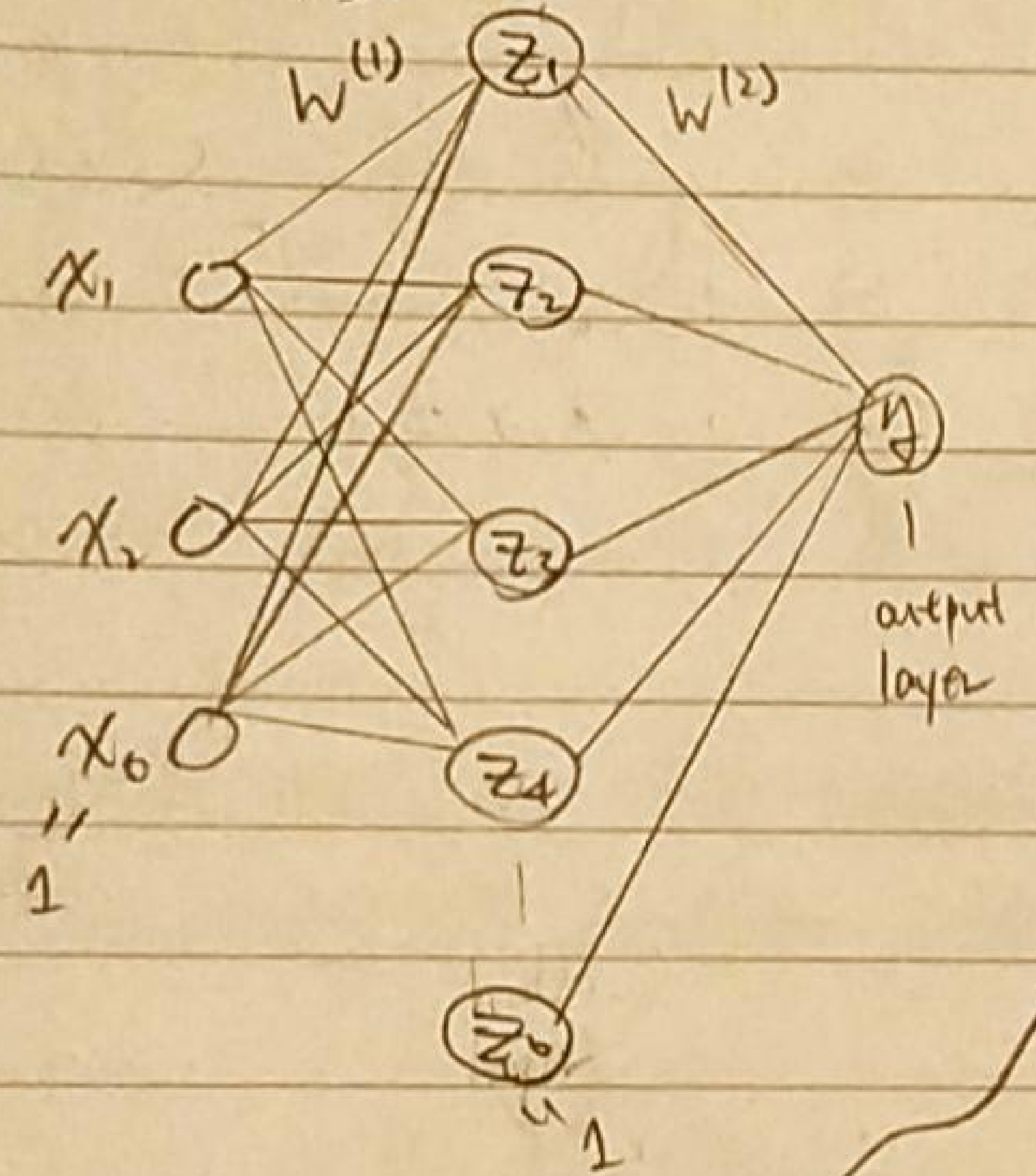
Assignment - 2

Theoratical exercise

1)





We embed an extra element in our input space $X = [x_1, x_2]$

$$\tilde{x} = \underbrace{[x_1, x_2, 1]^T}_{\text{biase}}$$

then, we define our hidden layer

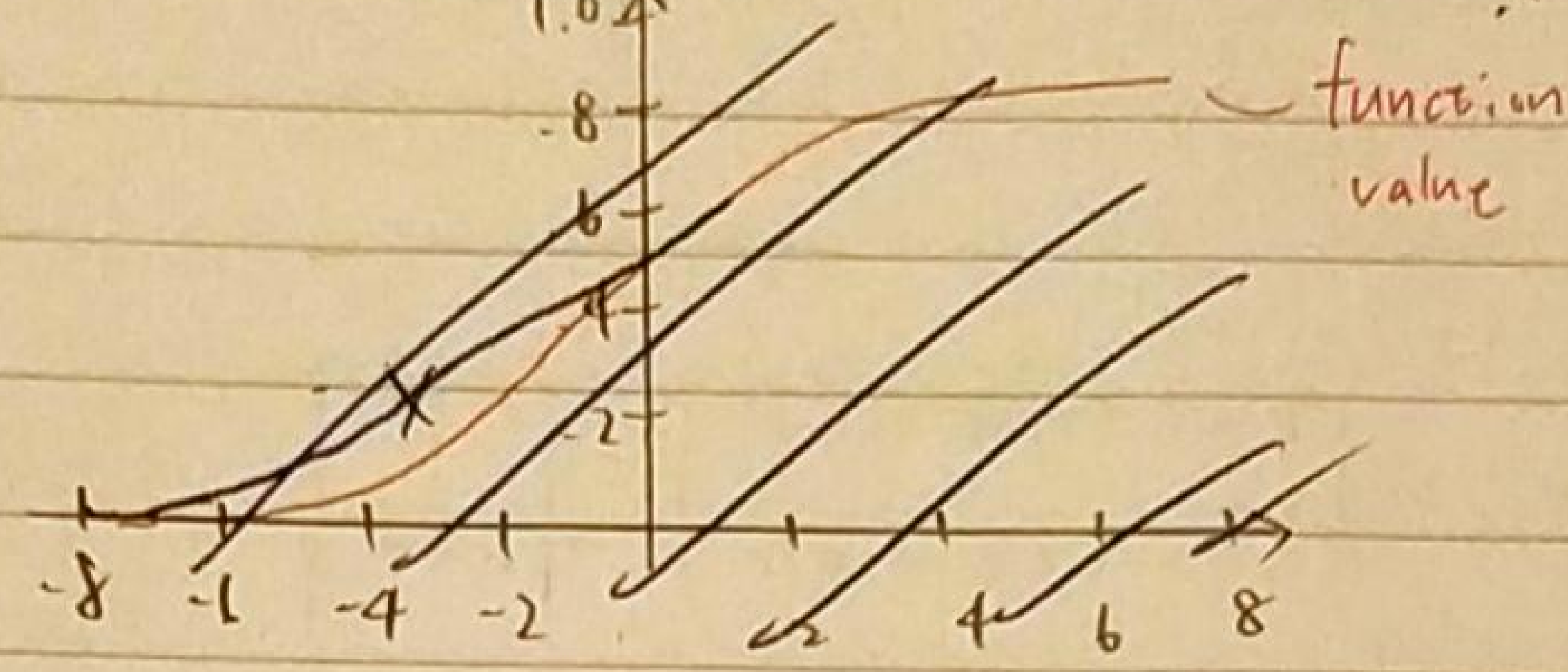
$$\boxed{\begin{aligned} z_1 &= h(a_1) = W_1^T \tilde{x}, \\ z_2 &= h(a_2) = W_2^T \tilde{x} \\ z_3 &= h(a_3) = W_3^T \tilde{x} \\ z_4 &= h(a_4) = W_4^T \tilde{x} \end{aligned}}$$

We can put all these in a matrix representation

$$A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = W^{(1)T} X = \begin{bmatrix} 1 & -2 & 6 \\ -1 & 3 & 2 \\ -1 & -3 & 14 \\ 2 & -1 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$

Now we define our activation function $h(\cdot)$

$$z(a) = h(a) = \sigma_{sigmoid}(a) = \frac{1}{1+\exp(-a)}$$



all the points inside the trapezoid will lead to a positive value of $z(a)$ and sum of z_1, z_2, z_3, z_4 will be 4 or close to 4, since we are using sigmoid as activation function

For the output layer, we define

$y = \sigma_{sigmoid}(W^{(2)T} Z)$, where $W^{(2)}$ and Z are

$$W^{(2)} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -4 \end{bmatrix}, \quad Z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_0 \end{bmatrix}$$

weight   bias

the -4 in layer 2 allow the final output neuron act as an AND Gate, only when all 4 neurons in hidden layer are active

output positive value (Probability)

Assignmen-2

2.

$$X = \begin{bmatrix} 0.1 \\ 0.4 \\ 1 \end{bmatrix} \qquad W^{(1)} = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.2 & 0.3 & 0.3 \end{bmatrix}$$

$$A^{(1)} = W^{(1)} X = \begin{bmatrix} 0.01 + 0.08 + 0.3 \\ 0.02 + 0.12 + 0.3 \end{bmatrix} = \begin{bmatrix} 0.39 \\ 0.44 \end{bmatrix}$$

$$h(A^{(1)}) = \begin{bmatrix} \sigma_s(0.39) \\ \sigma_s(0.44) \end{bmatrix} = \begin{bmatrix} 0.59 \\ 0.60 \end{bmatrix} \approx \begin{bmatrix} 0.60 \\ 0.60 \end{bmatrix}$$

$$\parallel$$

$$Z^{(1)}$$

$$Z^{(1)} = \begin{bmatrix} 0.6 \\ 0.6 \\ 1 \end{bmatrix} \qquad W^{(2)} = \begin{bmatrix} 0.4 & 0.5 & 0.6 \\ 0.5 & 0.6 & 0.6 \end{bmatrix}$$

$$A^{(2)} = W^{(2)} Z^{(1)} = \begin{bmatrix} 0.24 + 0.3 + 0.36 \\ 0.3 + 0.36 + 0.6 \end{bmatrix} = \begin{bmatrix} 0.9 \\ 1.26 \end{bmatrix}$$

$$O = h(A^{(2)}) = \begin{bmatrix} 0.71 \\ 0.77 \end{bmatrix}$$

$$E(W, x) = \frac{1}{2} \| O - t \|^2$$

$$= \frac{1}{2} \left\| \begin{bmatrix} 0.71 \\ 0.77 \end{bmatrix} - \begin{bmatrix} 0.1 \\ 0.9 \end{bmatrix} \right\|^2$$

$$= \frac{1}{2} \left[(0.61)^2 + (-0.13)^2 \right]$$

$$= \underline{0.1945}$$