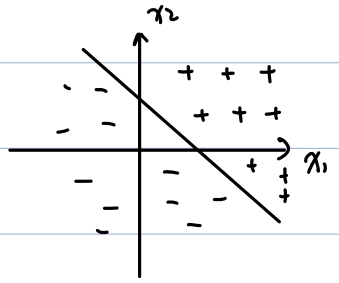
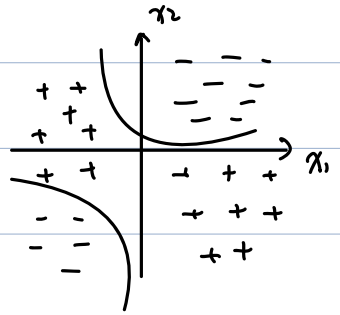
Lec 13 Multi-layer perceptron MLP

1. Recall:

Perceptron: $\hat{y} = \text{sign}(w^T x)$



linearly sepertable



Not linearly sepertable

⇒ solve by { pocket algo.
Non-linear transformation

↳ e.g. $Z = \{1, x_1, x_2, x_1^2, x_2^2\}$

problem: need to know what transform to use ahead of time

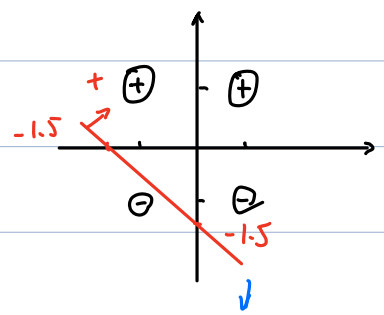
↳ what to do ? ⇒ logic gates.

2. Logic gates:

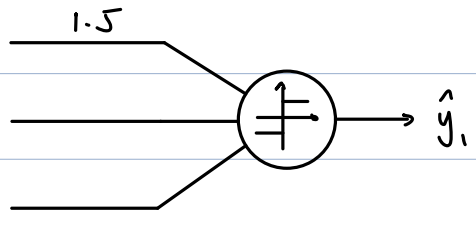
↳ input: $h_1, h_2 \in \{+1, -1\}$

↳ $y_1 = OR(h_1, h_2)$ $y_2 = AND(h_1, h_2)$ $y_3 = XOR(h_1, h_2)$

↳. OR (h_1, h_2):



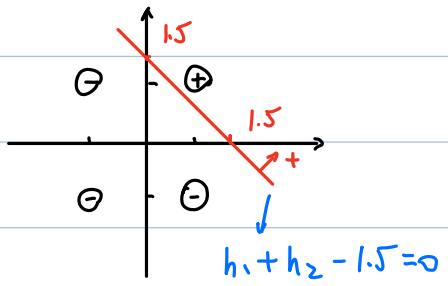
$\therefore$



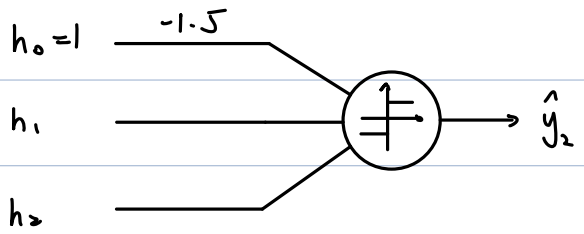
$\begin{cases} h_1 + h_2 + 1.5 = 0 \checkmark \\ -h_1 - h_2 - 1.5 = 0 \times \end{cases}$

// same line. but w is different

↳ AND (h_1, h_2)

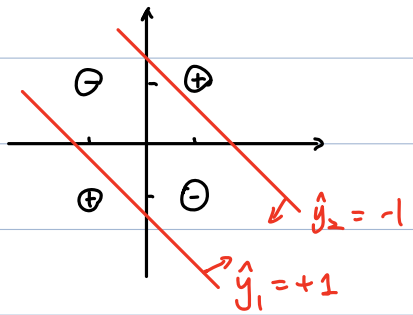


∴



$\therefore \hat{y}_2 = \text{sign}(h_1 + h_2 - 1.5)$

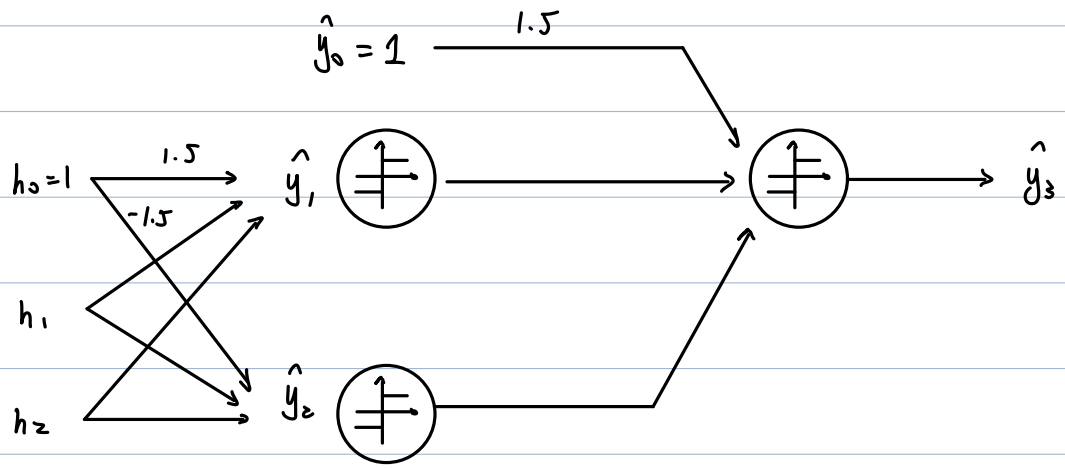
↳ XOR (h_1, h_2)



idea: express XOR using AND, OR

e.g. $\hat{y}_3 = \text{AND}(\hat{y}_1, \bar{\hat{y}}_2)$

// XOR = OR but not AND

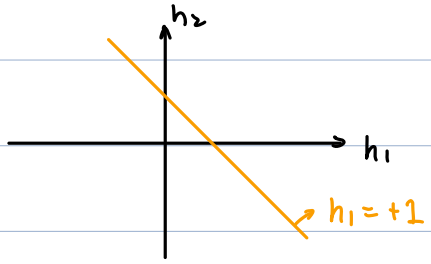


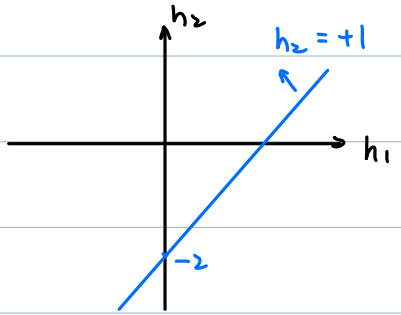
↑

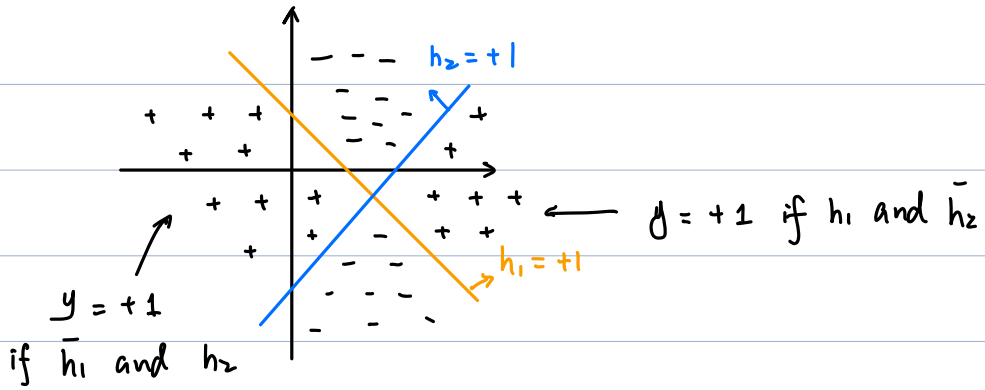
Multi-layer perceptron

3. Need one more gate to convert $f(x)$ into h_1, h_2

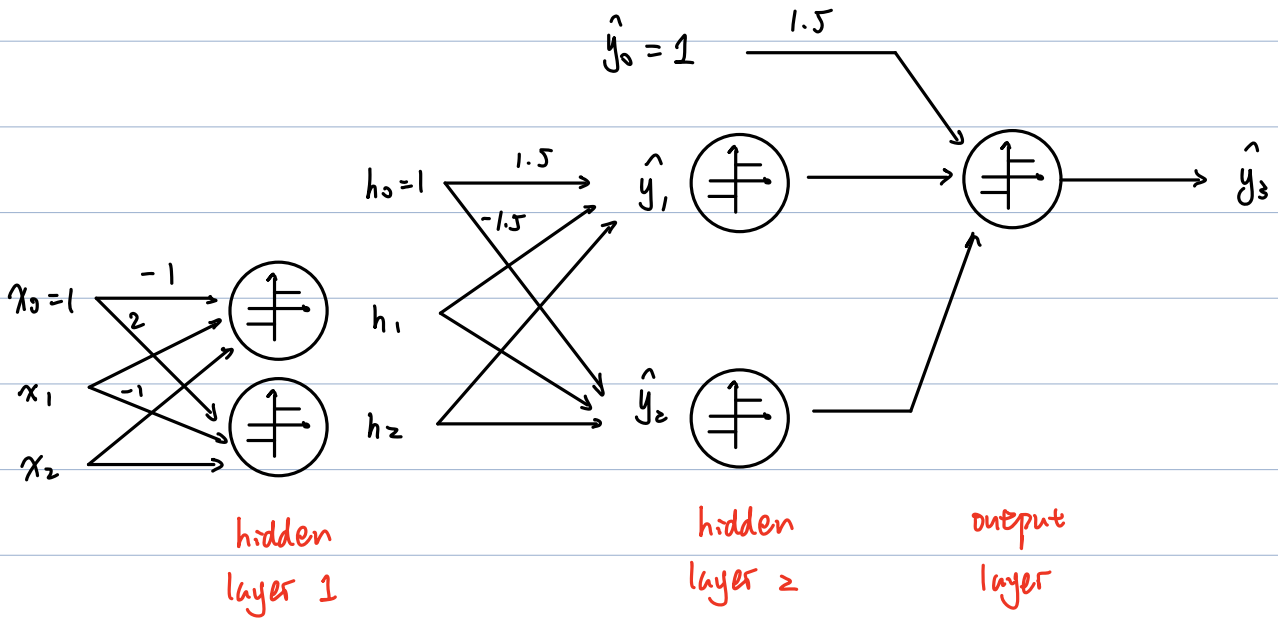
e.g. $h_1 = \text{sign}(x_1 + x_2 - 1)$ $\quad h_2 = \text{sign}(-x_1 + x_2 + 2)$







More precisely, $y = \text{XOR}(h_1, h_2)$



3-layer MLP