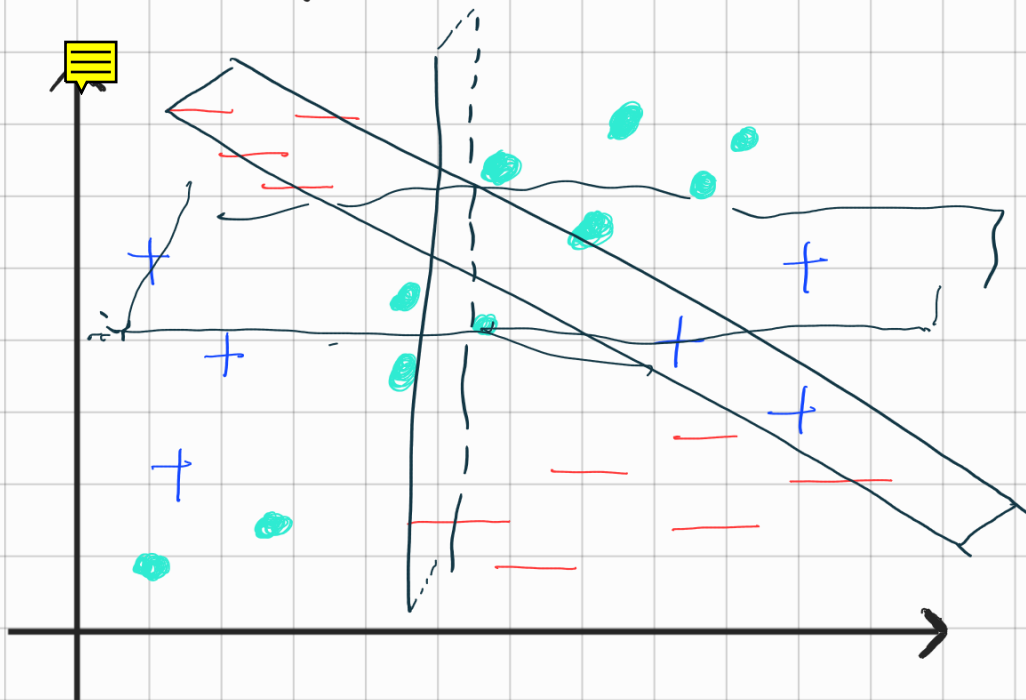
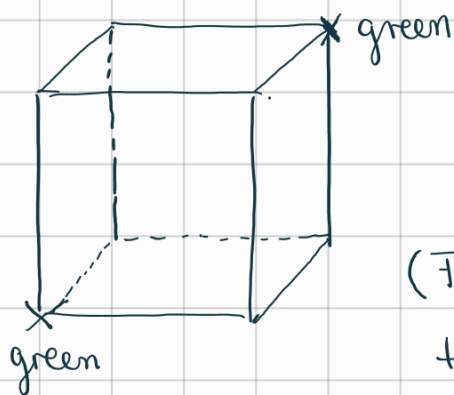


1. Hand-Crafted Network



The general idea in this case is to use two corners of the $\{0,1\}^3$ hypercube for one of the classes present here. In this case:



(Fig.1: For instance green takes these two corners)

The first layer of the neural network classifies the data as the depicted Y_i in the specific region of the corner.

Class 1 $\rightarrow \{ (1,0,0), (0,1,1) \}$

Class 2 $\rightarrow \{ (0,1,0), (1,0,1) \}$

Class 3 $\rightarrow \{ (0,0,1), (1,1,0) \}$

2. Linear Activation Function

$$z_0 = x$$



$$\bar{z}_l = z_{l-1} \cdot B_l + b_l$$

$$z_l = \phi_l(\bar{z}_l)$$

If ϕ_l is the identity function then any network with depth $L > 1$ is equivalent to a 1-layer neural network.

Proof:

$$\begin{aligned} \text{If } \phi_l \text{ is the identity then} \\ z_l &= \phi_l(\bar{z}_l) \\ &= \text{id}(\bar{z}_l) \\ &= \text{id}(z_{l-1} \cdot B_l + b_l) \\ &= z_{l-1} \cdot B_l + b_l \end{aligned}$$

$$\begin{aligned} z_3 &= z_2 \cdot B_3 + b_3 \\ &= (z_1 \cdot B_2 + b_2) \cdot B_3 + b_3 \\ &= ((z_0 \cdot B_1 + b_1) \cdot B_2 + b_2) \cdot B_3 + b_3 \\ &= (x B_1 B_2 + b_1 B_2 + b_2) \cdot B_3 + b_3 \\ &= x B_1 B_2 B_3 + b_1 B_2 B_3 + b_2 B_3 + b_3 \end{aligned}$$

Let l be greater than 1

$$b_1 B_2 B_3 + b_2 B_3 + b_3$$

$$\begin{aligned} z_l &= z_{l-1} \cdot B_l + b_l \\ &= x \cdot \prod_{i=1}^l B_i + \sum_{i=1}^l b_i \prod_{j=i+1}^{l-1} B_{j+1} \end{aligned}$$

$$\text{Now we define } B := \prod_{i=1}^l B_i \text{ and } b := \sum_{i=1}^l b_i \prod_{j=i+1}^{l-1} B_{j+1}$$

$$z_l = x \cdot B + b$$

$$\begin{aligned} \text{This is the same expression as } z_1 &= z_0 \cdot B_1 + b_1 \\ &= x \cdot B_1 + b_1 \end{aligned}$$

So in this case any network with depth $L > 1$ is equivalent to a 1-layer neural network

