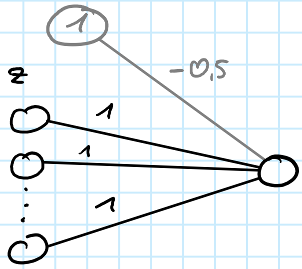
# Task 1, 2

Donnerstag, 6. Mai 2021 14:56

Lars Kühmichel
Nicolas Wolf

## 1.1

a) $z \in \{0,1\}^n \mapsto f(z) = \begin{cases} 0 & \text{if } z_i = 0 \ \forall i \\ 1 & \text{otherwise} \end{cases}$



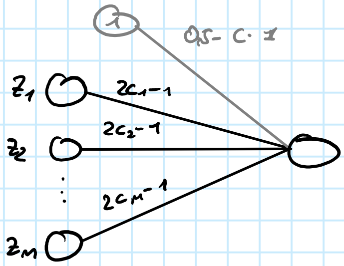

$$\Rightarrow \tilde{z} = (1, z) \cdot \begin{pmatrix} -0{,}5 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \underbrace{\sum_i z_i}_{=0 \text{ iff } z=0} - 0{,}5$$

$$f(z) = \Theta(\tilde{z}) = \begin{cases} 0 & \text{if } \tilde{z} \leq 0 \\ 1 & \text{if } \tilde{z} > 0 \end{cases}$$

b) let $c \in \{0,1\}^m$ be fixed, $z \in \{0,1\}^m \mapsto f(z) = \begin{cases} 1 & z = c \\ 0 & \text{otherwise} \end{cases}$

$$\beta_i = \begin{cases} 1 & \text{if } c_i = 1 \\ -1 & \text{if } c_i = 0 \end{cases} = 2c - \mathbb{1} \qquad b = 0{,}5 - c.\text{sum}()$$

activation function: $\varphi = \Theta$




$$\tilde{z} = z \cdot \beta + b = z \cdot (2c - 1) + b$$
$$= 2\, z \cdot c - z \cdot \mathbb{1} - c \cdot \mathbb{1} + 0{,}5$$
$$= 2 \cdot z \cdot c - (z + c) \cdot \mathbb{1} + 0{,}5$$

- if $z = c$: $z \cdot c = c \cdot c = \sum c_i^2 = \sum c_i \Rightarrow \tilde{z} = 0{,}5$
- if $z \neq c$: $\exists i: z_i \cdot c_i = 0$

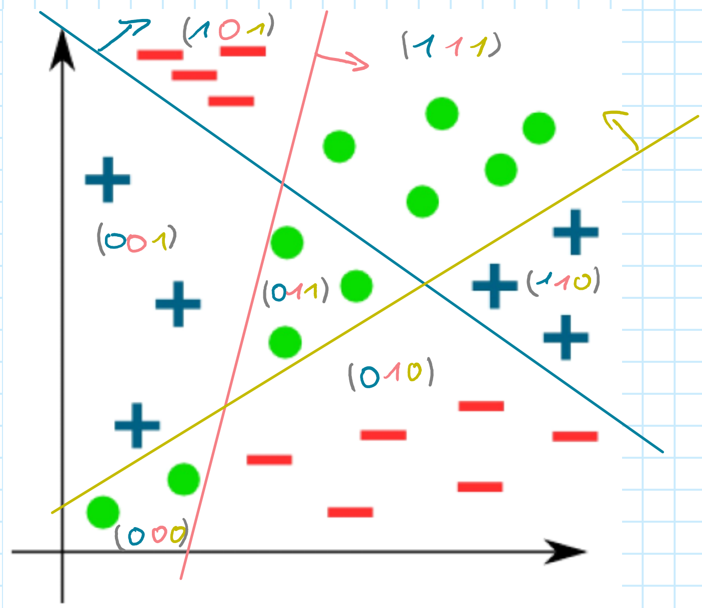
$$\Rightarrow z \cdot c = \sum_i z_i \cdot c_i < \sum_i c_i^2 = \sum_i c_i = c \cdot \mathbb{1}$$

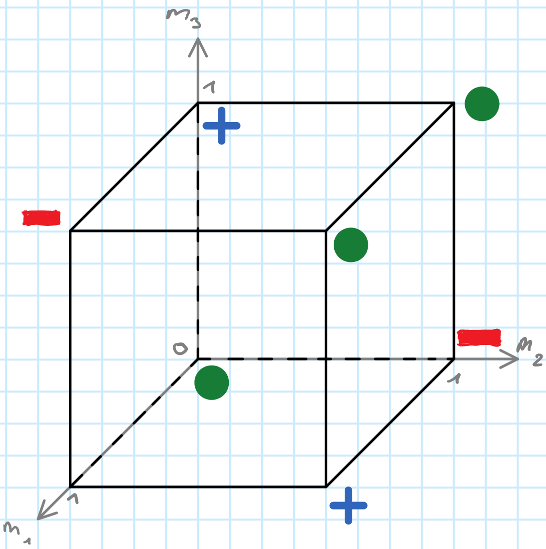
and $\sum_i z_i \cdot c_i < \sum_i z_i^2 = \sum_i z_i = z \cdot \mathbb{1}$

$$\Rightarrow 2 \cdot z \cdot c < (z + c) \cdot \mathbb{1}$$

$$\Rightarrow \tilde{z} < 0$$

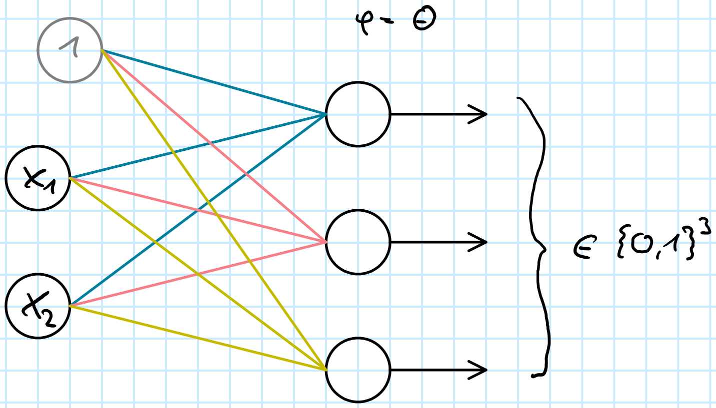
c)



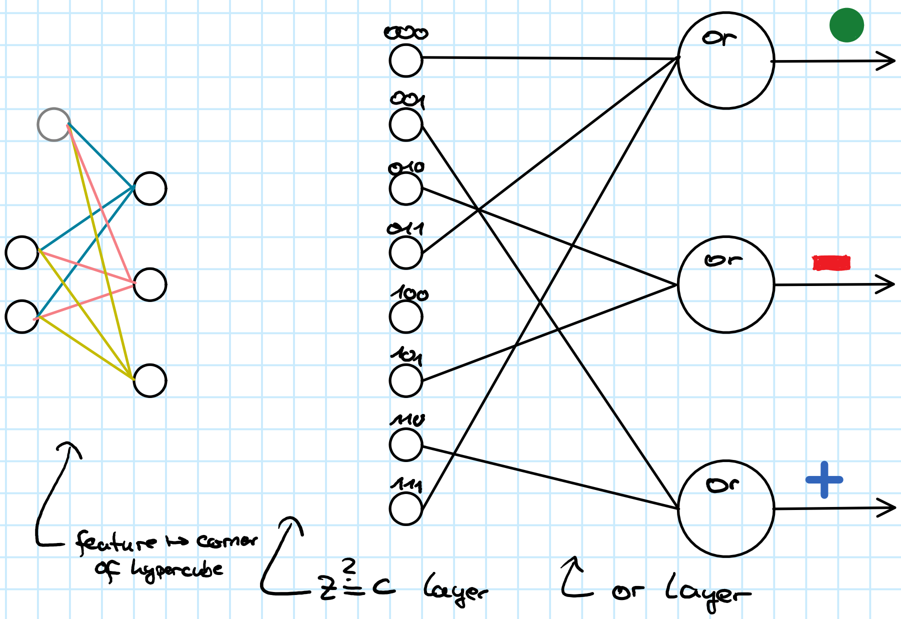




$\varphi = \Theta$

1




1.3 • 000, 011, 111 — 010, 101 + 001, 110




- First, we map the features onto the corners of $\{0,1\}^m$ (for the given dataset: $m=3$). Since each corner contains maximal 1 class, we can use the network from 1.1.2 to test, if the output equals $c \in \{0,1\}^m$, which we do for every combination. Only one of those $2^m$ nodes will output a 1. Since we know the class labels of the hypercube corners, we combine the outputs using the or-network, s.t. one or-node is linked to all $z \stackrel{?}{=} c$-nodes with the same class label (the weights to the other $z \stackrel{?}{=} c$-nodes are set to 0).

- For this example, we only need a 3-dimensional hypercube, but for other data sets $m$ could be larger which may lead to overfitting

② • Lemma: Let $f: \mathbb{R}^l \to \mathbb{R}^m$, $g: \mathbb{R}^n \to \mathbb{R}^l$ linear. Then $f \circ g: \mathbb{R}^n \to \mathbb{R}^m$ is linear

Prove:

2. • Lemma: Let $f: \mathbb{R}^l \to \mathbb{R}^m$, $g: \mathbb{R}^n \to \mathbb{R}^l$ Linear. Then $f \circ g: \mathbb{R}^n \to \mathbb{R}^m$ is Linear
Prove: Let $v, \omega \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$

$$(f \circ g)(\lambda v + \omega) = f(g(\lambda v + \omega)) \overset{g \text{ Linear}}{=} f(\lambda g(v) + g(\omega)) \overset{f \text{ Linear}}{=} \lambda f(g(v)) + f(g(\omega))$$

• If $\varphi_l$ is Linear, then $z_{l-1} \mapsto z_l$ is Linear

$$z_l = \varphi_l(\tilde{z}_l) = \varphi_l\left(\underbrace{(1, z_{l-1}) \cdot \begin{pmatrix} b_l \\ B_l \end{pmatrix}}_{\text{Linear}}\right)$$

• Induction: • $z_1(z_0)$ is Linear
  • assume $z_l(z_0)$ is Linear
  • $z_{l+1} = \underset{\text{Linear}}{\varphi_{l+1}}\left(\underbrace{(1, \underbrace{z_l(z_0)}_{\text{Linear}}) \cdot \begin{pmatrix} b_{l+1} \\ B_{l+1} \end{pmatrix}}_{\text{Linear}}\right) \Rightarrow z_{l+1}(z_0)$ Linear

• $\Rightarrow Y = z_L(X)$ is Linear
  $\Rightarrow$ can find transformation matrix $B$, s.t. $Y = X \cdot B$
  $\Rightarrow$ use $B$ as weights for 1-Layer network