Sonntag, 19. Mai 2024 16:11

1) (1)

$z_0 = 1$

$\beta_0 - b = -1/2$

$\beta = (-1/2, \underbrace{1, 1, \dots, 1}_{D})$

$f = \frac{1}{2}(1 + \text{sign}(z \cdot \beta))$

(2) $f = \frac{1}{2}(1 + \text{sign}(z_0 b + c \cdot z))$

$c_0 = -\frac{1}{2}$

$\beta = (b, \bar{c})$

(3) $h(c, z) = \begin{cases} 1 & z = c \\ 0 & \text{else} \end{cases}$

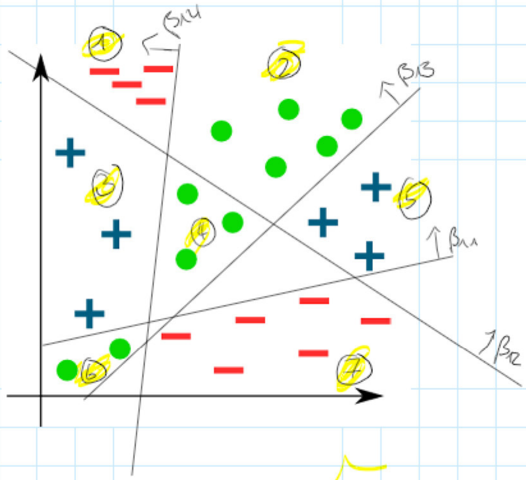
$f = \text{sign}(1 + \text{sign}(\underbrace{z \cdot \beta}_{\geq 0}))$

$-z^2$

$b = -z^2$

$\beta = (b, z)$

$\Rightarrow f = \begin{cases} -z^2 + z^2 = 0 & \text{if } z = c \\ \leq 0 & \text{if } z \neq c \end{cases}$

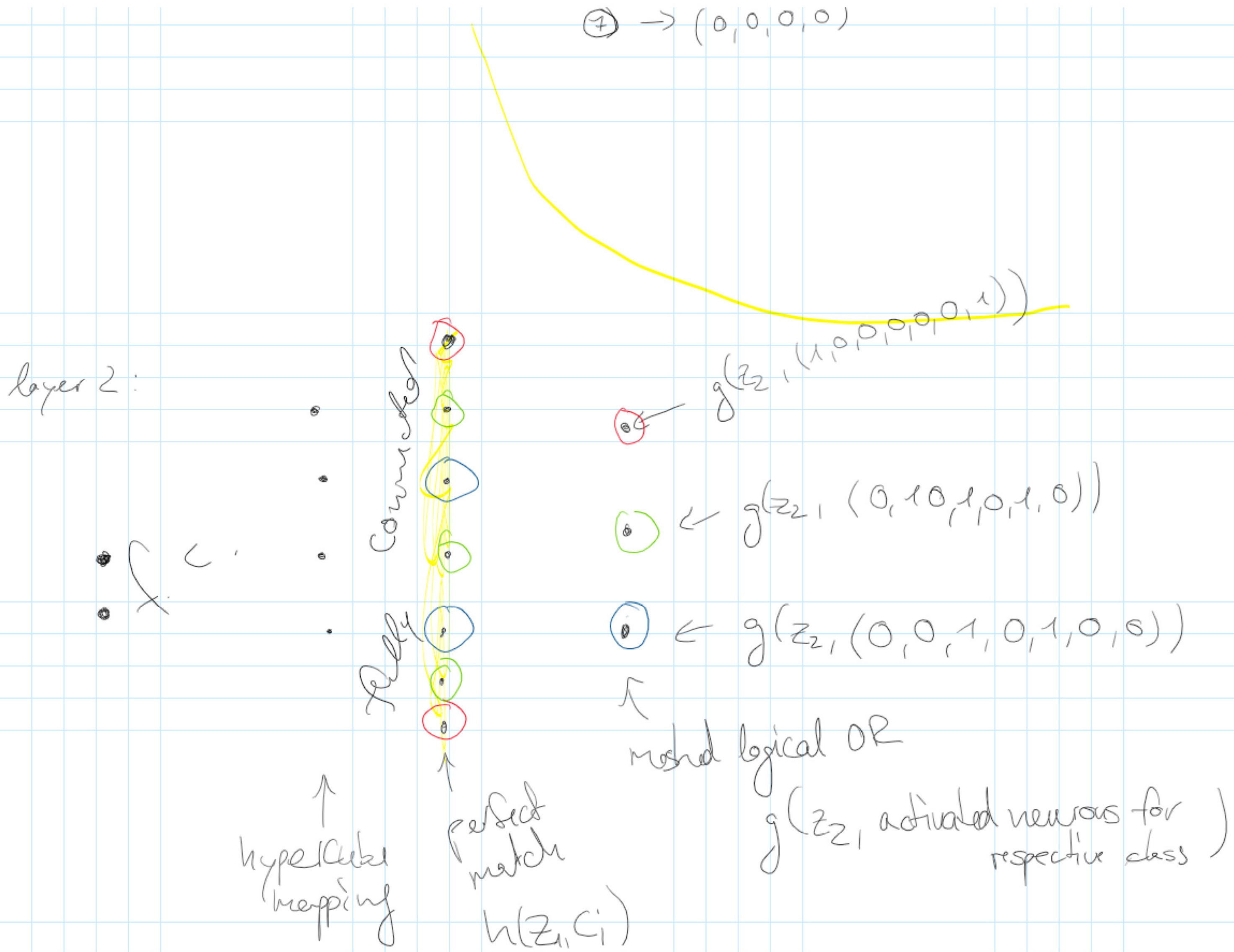


general corner coord

$\left(\frac{1}{2}(1 + \text{sign}(\beta_{11})), \frac{1}{2}(1 + \text{sign}(\beta_{12})), \frac{1}{2}(1 + \text{sign}(\beta_{13})), \frac{1}{2}(1 + \text{sign}(\beta_{14}))\right)$

$$\left.\begin{array}{l}
1 \to (1, 1, 1, 1) \\
2 \to (1, 1, 1, 0) \\
3 \to (1, 0, 1, 1) \\
4 \to (1, 0, 1, 0) \\
5 \to (1, 1, 0, 0) \\
6 \to (0, 0, 1, 1)
\end{array}\right\} C_i$$

$7 \to (0, 0, 0, 0)$



2 Linear Activation Function (5 Points)

For a feed forward network the output of each layer l is calculated iteratively by

$$
\begin{aligned}
Z_0 &= X && (1)\\
\tilde{Z}_l &= Z_{l-1} \cdot B_l + b_l \quad \text{(multiply with weights and add bias vector)} && (2)\\
Z_l &= \phi_l(\tilde{Z}_l) \quad \text{(apply the activation function element-wise to pre-activations)} && (3)
\end{aligned}
$$

Where X and the Z_l and b_l are understood as row vectors.

Prove that if ϕ_l is the identity function then any network (with depth $L > 1$) is equivalent to a 1-layer neural network.

Assumption $\tilde{z}_\ell = z_\ell \Rightarrow$

für Tiefe $L-1$ kann alles als Network mit Tiefe $L=1$ geschrieben werden

$$\tilde{z}_\ell = z_{\ell-1} \cdot B_\ell + b_\ell = \underbrace{(\underbrace{z_{\ell-2} \cdot B_{\ell-1} + b_{\ell-1}}_{\text{linear operation}}) B_\ell}_{\text{linear operation}} + b_\ell$$

Verkettung von lin. op. => lin. op. , ausdrückbar in einem.