Q1) 1)

$$\begin{bmatrix} w_{00} & w_{01} \\ w_{10} & w_{11} \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & x_{00} & x_{01} & x_{02} & 0 \\ 0 & x_{10} & x_{11} & x_{12} & 0 \\ 0 & x_{20} & x_{21} & x_{22} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Stride 3, padding 1.

$Y \in \mathbb{R}^{4 \times 1}$ $\quad X \in \mathbb{R}^{9 \times 1}$ $\quad \Rightarrow A \in \mathbb{R}^{4 \times 9}$

$$\boxed{A = \overset{4 \times 9}{\begin{bmatrix} w_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & w_{10} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & w_{01} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{00} \end{bmatrix}}} \quad \overset{9 \times 1}{\begin{bmatrix} x_{00} \\ x_{01} \\ x_{02} \\ x_{10} \\ x_{11} \\ x_{12} \\ x_{20} \\ x_{21} \\ x_{22} \end{bmatrix}}$$

Q1) 2)

$$W = \begin{bmatrix} w_{00} & w_{01} \\ w_{10} & w_{11} \end{bmatrix} \qquad x = \begin{bmatrix} x_{00} & x_{01} \\ x_{10} & x_{11} \end{bmatrix}$$

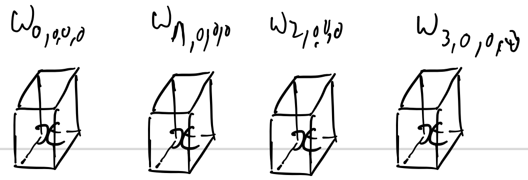
Stride 2, no padding.

$$A = \boxed{\begin{bmatrix} w_{00} & 0 & 0 & 0 \\ w_{01} & 0 & 0 & 0 \\ 0 & w_{00} & 0 & 0 \\ 0 & w_{01} & 0 & 0 \\ w_{10} & 0 & 0 & 0 \\ w_{11} & 0 & 0 & 0 \\ 0 & w_{10} & 0 & 0 \\ 0 & w_{11} & 0 & 0 \\ 0 & 0 & w_{00} & 0 \\ 0 & 0 & w_{01} & 0 \\ 0 & 0 & 0 & w_{00} \\ 0 & 0 & 0 & w_{01} \\ 0 & 0 & w_{10} & 0 \\ 0 & 0 & w_{11} & 0 \\ 0 & 0 & 0 & w_{10} \\ 0 & 0 & 0 & w_{11} \end{bmatrix} \begin{bmatrix} x_{00} \\ x_{01} \\ x_{10} \\ x_{11} \end{bmatrix}}$$

Generated y:

$$\begin{bmatrix} w_{00}x_{00} & w_{01}x_{00} & w_{00}x_{01} & w_{01}x_{01} \\ w_{10}x_{00} & w_{11}x_{00} & w_{10}x_{01} & w_{11}x_{01} \\ w_{00}x_{10} & w_{01}x_{10} & w_{00}x_{11} & w_{01}x_{11} \\ w_{10}x_{10} & w_{11}x_{10} & w_{10}x_{11} & w_{11}x_{11} \end{bmatrix}$$

Q1) 3)



$$\begin{bmatrix} x_{00} & x_{01} & x_{10} & x_{11} \end{bmatrix}$$

$$\text{size}(Y) = \begin{bmatrix} 4 & 4 & 1 \end{bmatrix}$$

$$\Rightarrow Y = \begin{bmatrix} w_{0000}x_{00} & w_{0000}x_{01} & w_{0000}x_{10} & w_{0000}x_{11} \end{bmatrix}$$

$$\begin{bmatrix} w_{1000}x_{00} & w_{1000}x_{01} & w_{1000}x_{10} & w_{1000}x_{11} \end{bmatrix}$$

w_{2000} ...

w_{3000} ...

Affine Transformation:

size 4×4 size 4×4 size 4×4 size 4×4

$$\Rightarrow W_{0,0} = \text{diag}(w_{0000}) \qquad W_{1,0} = \text{diag}(w_{1000}) \qquad W_{2,0} = \text{diag}(w_{2000}) \qquad W_{3,0} = \text{diag}(w_{3000})$$

$$W \in \mathbb{R}^{4\times1\times4\times4}$$

$$\Rightarrow \text{size}(Y_t) = \begin{bmatrix} 1 & 4 & 4 \end{bmatrix} \qquad \begin{bmatrix} x_{00} & x_{01} & x_{10} & x_{11} \end{bmatrix}$$

$$Y_t = \begin{bmatrix} w_{00}x_{00} & w_{01}x_{00} & w_{00}x_{01} & w_{01}x_{01} \\ w_{10}x_{00} & w_{11}x_{00} & w_{10}x_{01} & w_{11}x_{01} \\ w_{00}x_{10} & w_{01}x_{10} & w_{00}x_{11} & w_{01}x_{11} \\ w_{10}x_{10} & w_{11}x_{10} & w_{10}x_{11} & w_{11}x_{11} \end{bmatrix} \Rightarrow$$

$\exists$ a bijection from the indices of Y to Y_t to find the corresponding elements that are equal.

ng

$$A = \begin{bmatrix} w_{00} & 0 & 0 & 0 \\ w_{01} & 0 & 0 & 0 \\ 0 & w_{00} & 0 & 0 \\ 0 & w_{01} & 0 & 0 \\ w_{10} & 0 & 0 & 0 \\ w_{11} & 0 & 0 & 0 \\ 0 & w_{10} & 0 & 0 \\ 0 & w_{11} & 0 & 0 \\ 0 & 0 & w_{00} & 0 \\ 0 & 0 & w_{01} & 0 \\ 0 & 0 & 0 & w_{00} \\ 0 & 0 & 0 & w_{01} \\ 0 & 0 & w_{10} & 0 \\ 0 & 0 & w_{11} & 0 \\ 0 & 0 & 0 & w_{10} \\ 0 & 0 & 0 & w_{11} \end{bmatrix}$$

Q2) 1)

AND: $w_{AND} = [1 \quad 1]^T$

$b_{AND} = -2$

$x = [0 \quad 1] \rightarrow \quad w^T x + b = 1 - 2 = -1 < 0 \Rightarrow f(x) = 0$

$x = [1 \quad 1] \rightarrow \quad w^T x + b = 2 - 2 = 0 \geq 0 \Rightarrow f(x) = 1$ ✓

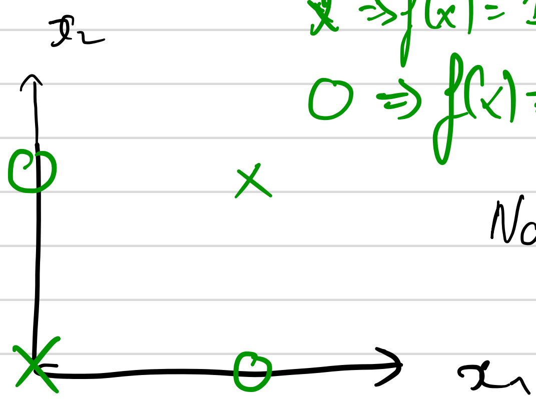
OR: $w_{OR} = [1 \quad 1]^T$

$b_{OR} = -1$

$x = [0 \; 1] \rightarrow \quad w^T x + b = 1 - 1 \geq 0 \Rightarrow f(x) = 1$

$x = [0 \; 0] \rightarrow \quad w^T x + b = 0 - 1 < 0 \Rightarrow f(x) = 0$ ✓

2)



$\times \Rightarrow f(x) = 1$

$O \Rightarrow f(x) = 0$

No line can seperate the Xs & Os.

More formally:

$[1\ 0] \Rightarrow f(x) = 1: \quad w_1 + b \geq 0 \quad ①$

$[0\ 1] \Rightarrow f(x) = 1: \quad w_2 + b \geq 0 \quad ②$

$[1\ 1] \Rightarrow f(x) = 0: \quad w_1 + w_2 + b < 0 \quad ③$

$[0\ 0] \Rightarrow f(x) = 0: \quad b < 0 \quad ④$

$① \,\&\, ④ \Rightarrow w_1 \overset{⑤}{\geq} -b > 0 \qquad ② \,\&\, ④ \Rightarrow w_2 \overset{⑥}{\geq} -b > 0$:

Plug ⑤ & ⑥ into ③ $\Rightarrow \quad -b - b + b < 0 \Rightarrow -b < 0 \Rightarrow b > 0$,

but $b < 0$!!!

So impossible to satisfy all ④ equations.

Q3)

1) $x = 1 \Rightarrow h(x) = 5$, $\{W = 4 , b = 1\} \Rightarrow Wx + b = 4 + 1 = 5$ ∅

$$\left.\frac{dh}{dx}\right|_{x=1} = W^{(3)} W^{(2)} W^{(1)} = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} = 2$$

Q3)

2) $x = -1 \Rightarrow h(x) = 2$, $\{W = -1, b = 1\} \Rightarrow Wx + b = 1 + 1 = 2$ $\emptyset$

$$\left.\frac{dh}{dx}\right|_{x=-1} = \begin{array}{c} w_s \\ \begin{bmatrix} 1 & 1 \end{bmatrix} \end{array} \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0.5 \end{bmatrix} = 1$$

Q3)

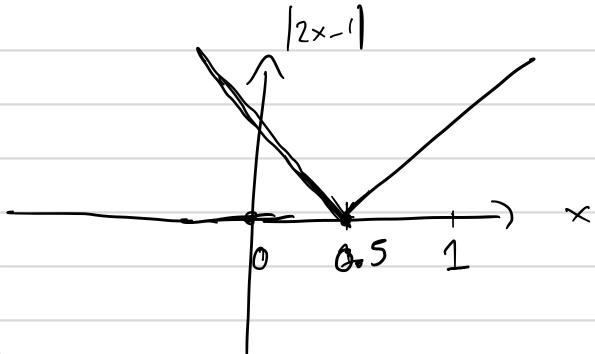
3) $x = -0.5 \Rightarrow h(x) = 2.5$, $\{W = -3 , b = 1\} \Rightarrow Wx + b = 1.5 + 1 = 2.5$

$$\boxed{\left.\frac{dh}{dx}\right|_{x=1} = W^{(3)} W^{(2)} W^{(1)} = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0.5 \end{bmatrix} = 1}$$

Q4) i)

$$\begin{array}{ccccc} d\times d & & d\times 1 & & \\ \begin{bmatrix} 2 & & & & \\ & 2 & & 0 & \\ 0 & & 2 & & \\ & & & 2 & \\ & & & & \ddots \end{bmatrix} & & \begin{bmatrix} \\ \\ \\ \\ \\ \end{bmatrix} & - & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{array} = 2x - 1$$

For $d = 1$:



$x \in \{0, 0.5\}$ & $x \in \{0.5, 1\}$

So for each d, 2 regions are identified.

For $d = 2$, how many region combinations are possible?

$x_1 \in \{0, 0.5\}$, $x_1 \in \{0.5, 1\}$

$x_2 \in \{0, 0.5\}$, $x_2 \in \{0.5, 1\}$

Q4) 2) How many regions of its input does $f \circ g(\cdot)$ identify onto $(0,1)^d$?

We know $f(\cdot)$ identifies 2^d regions of its inputs onto $(0,1)^d$. For each one of those regions, $g(\cdot)$ can identify 2^d regions of its input onto it.

Hence $\boxed{n_{fg} = n_f \, n_g} = 2^{2d}$

Q4) 3) Following the logic of the previous questions, each layer can identify 2^d regions of its inputs to one of the desired regions of the following layer. Starting from the last layer we get,

$$\boxed{n = \underbrace{2^d \cdot 2^d \cdot 2^d \cdot 2^d \cdots}_{L\text{-Times}} = 2^{Ld}}$$

Last layer