CS 4803 | PS2

Reagan Kan | rkan3 | 903 404 746

1.1)

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

zero-pad size = 1

stride = 2

y shape = $(4,1)$

X shape = $(9,1)$

→ A shape = $(4,9)$

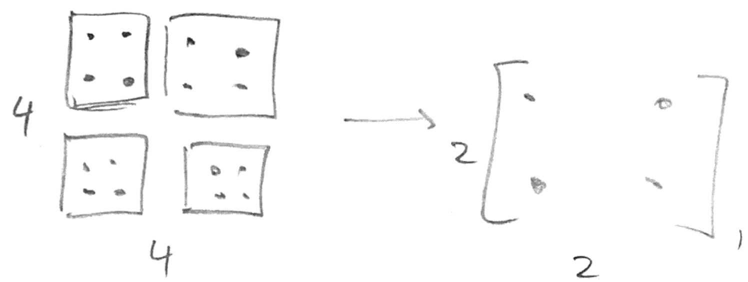
$$A = \begin{array}{ccccccccc} X_{00} & X_{01} & X_{02} & X_{10} & X_{11} & X_{12} & X_{20} & X_{21} & X_{22} \end{array}$$
$$\begin{bmatrix} W_{11} & W_{12} & 0 & W_{21} & W_{22} & 0 & 0 & 0 & 0 \\ 0 & W_{10} & W_{11} & 0 & W_{20} & W_{21} & 0 & 0 & 0 \\ 0 & 0 & 0 & W_{01} & W_{02} & 0 & W_{11} & W_{12} & 0 \\ 0 & 0 & 0 & 0 & W_{00} & W_{01} & 0 & W_{10} & W_{11} \end{bmatrix}$$

So, $y = A \cdot X$

$$\begin{bmatrix} y_{00} \\ y_{01} \\ y_{10} \\ y_{11} \end{bmatrix} = \begin{bmatrix} W_{11} & W_{12} & 0 & W_{21} & W_{22} & 0 & 0 & 0 & 0 \\ 0 & W_{10} & W_{11} & 0 & W_{20} & W_{21} & 0 & 0 & 0 \\ 0 & 0 & 0 & W_{01} & W_{02} & 0 & W_{11} & W_{12} & 0 \\ 0 & 0 & 0 & 0 & W_{00} & W_{01} & 0 & W_{10} & W_{11} \end{bmatrix} \cdot \begin{bmatrix} X_{00} \\ X_{01} \\ X_{02} \\ X_{10} \\ X_{11} \\ X_{12} \\ X_{20} \\ X_{21} \\ X_{22} \end{bmatrix}$$

$\boxed{1\text{-}2}$ $W_{shape} = (2,2)$ $X_{shape} = (2,2)$. stride 2, no pad

So a forward convolution takes (r, c), does stride 2 w/ 2x2 kernel, no pad, to get 2x2 output.



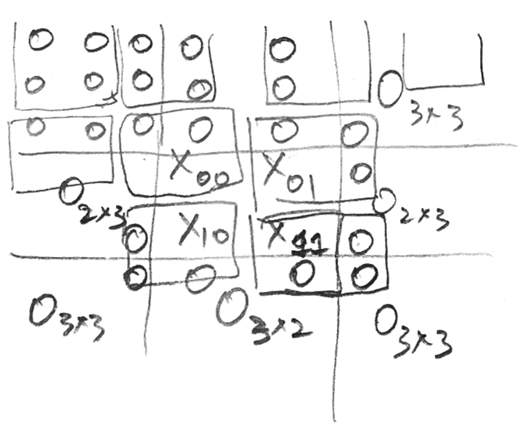
, so $Y_{shape} = (4,4)$.

Must use same kernel (2×2) and stride (2) to convert input dim $(2+2p', 2+2p')$ to $(4,4)$

$$4 = ((2+2p') - 2)/2 + 1 \quad \text{from } ((W+2P) - F)/S + 1$$

$3 = p'$ So zero pad X w/ 3.

$y =$ Convolve 2x2, w/ 2 stride on the following:



row major
$4\times4 \to 16\times1$

$$y = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & X_{00}W_{11} & X_{01}W_{10} & 0 \\ 0 & X_{10}W_{01} & X_{11}W_{00} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Now A has shape (16×4)

Need to find A s.t. $y_{row} = A \begin{bmatrix} X_{00} \\ X_{01} \\ X_{10} \\ X_{11} \end{bmatrix} \longrightarrow$

$$y = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \hline 0 \\ w_{11} x_{00} \\ w_{10} x_{01} \\ 0 \\ \hline 0 \\ w_{01} x_{10} \\ w_{00} x_{11} \\ 0 \\ \hline 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{array}{c} \begin{bmatrix} \multicolumn{4}{c}{0_{5\times4}} \\ w_{11} & 0 & 0 & 0 \\ 0 & w_{10} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & w_{01} & 0 \\ 0 & 0 & 0 & w_{00} \\ \multicolumn{4}{c}{0_{5\times4}} \end{bmatrix} \\ \begin{array}{cccc} x_{00} & x_{01} & x_{10} & x_{11} \end{array} \end{array} \begin{bmatrix} x_{00} \\ x_{01} \\ x_{10} \\ x_{11} \end{bmatrix}$$

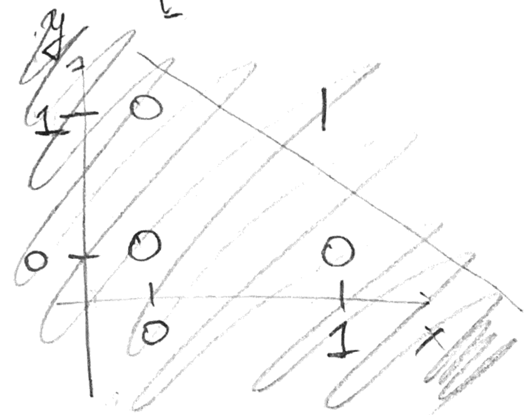
$$y = A \cdot X$$

2-1)

$w \in \mathbb{R}^2,\ b \in \mathbb{R},\ x \in \{0,1\}^2$

$$f(x) = \begin{cases} 1 & w^T x + b \geq 0 \longrightarrow w^T x \geq -b \\ 0 & w^T x + b < 0 \longrightarrow w^T x < -b \end{cases}$$

f_{AND}



$w_1(1) + w_2(1) \geq -b$ (a)

$w_1(0) + w_2(0) < -b$ (b)

$w_1(0) + w_2(1) < -b$ (c)

$w_1(1) + w_2(0) < -b$ (d)

from (b), $0 < -b$, so $b < 0$. Let's fix $b = -1$

then (a) becomes $w_1 + w_2 \geq 1$

(c) " $w_2 < 1$

(d) $w_1 < 1$

Letting $w_1 = w_2 = 0.6$ satisfies these 3 inequalities.

$$\boxed{W_{AND} = \begin{pmatrix} 0.6 \\ 0.6 \end{pmatrix} \quad b_{AND} = -1.0}$$

f_{OR}

$w_1(0) + w_2(0) < -b \rightarrow 0 < -b,\ b < 0,$ Again fix $b = -1$

$w_1(0) + w_2(1) \geq 1 \rightarrow w_2 \geq 1$ (a)

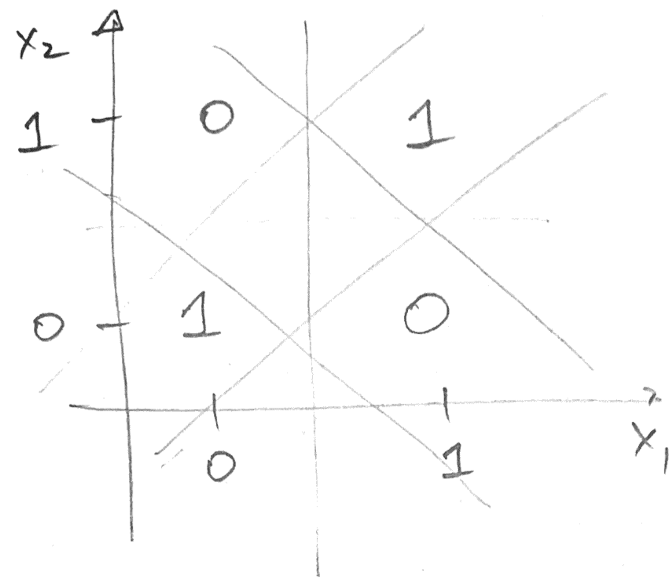
$w_1(1) + w_2(0) \geq 1 \rightarrow w_1 \geq 1$ (b)

$w_1(1) + w_2(1) \geq 1 \rightarrow w_1 + w_2 \geq 1$ (c)

Let $w_1 = w_2 = 1$, satisfies a, b, c.

$$\boxed{W_{OR} = \begin{pmatrix} 1.0 \\ 1.0 \end{pmatrix} \quad b_{OR} = -1.0}$$

2.2



$w_1(0) + w_2(0) \geq -b \rightarrow 0 \geq -b \rightarrow b \geq 0$ (a)

$w_1(1) + w_2(1) \geq -b \rightarrow w_1 + w_2 \geq -b$ ★

$w_1(1) + w_2(0) < -b \rightarrow w_1 < -b$ ★★

$w_1(0) + w_2(1) < -b \rightarrow w_2 < -b$ ★★★

from ★★, ★★★, $w_1 + w_2 < -2b$

$\Rightarrow w_1 + w_2 + 2b < 0.$ (b)

from ★, $w_1 + w_2 + b \geq 0$

Let $X = w_1 + w_2 + b$, then $X \geq 0.$ (c)

from (b), $(w_1 + w_2 + b) + b < 0, \Rightarrow X + b < 0$

from (a), (b), we know $X \geq 0$, $b \geq 0$, so $X + b \geq 0$.

this contradicts (b) $(X + b < 0)$

So "$<=>$" cannot be represented using linear model of given form.

3-1) $\sigma(\cdot) = f_1(\cdot) = \left| \vec{W}^{(1)} \vec{x} + \vec{b} \right| = \left| \underset{d\times d}{2 I} \cdot \underset{d\times 1}{\vec{x}} + \underset{d\times 1}{\vec{b}} \right|_{(d\times 1)}$

Look at each element in $|2I\vec{x}+\vec{b}|$

call it $|(2(1)x_i + b_i)|$ $\forall i \in \{1 \dots d\}$.

$b_i = -1$. (this is given).

We only care about $O = (0,1)^d$. open range

So $|2x_i - 1| < 1$.

$$-1 < 2x_i - 1 < 1$$

$$\frac{0}{2} < \frac{2x_i}{2} < \frac{2}{2}$$

$$0 < x_i < 1$$

So, for each element, $(0, 1)$ is the only input region that can be mapped to output region $(0,1)$ with given $W^{(1)}$, $b^{(1)}$.

each element of $\sigma(\cdot)$ has 1 input regions.

So, total of 1^d input regions $= \boxed{1 \text{ input region}}$

$\sigma(\cdot)$ is a bijection, since $\sigma^{-1}(\cdot)$ exists.

$$\sigma^{-1}(\cdot) = \frac{1}{2} I (\vec{x} - \vec{b})$$

3.2) $f \circ g(\cdot)$ identifies $n_g \cdot n_f$ regions onto $(0,1)^d$

3.3) from the explanation at the top of section 3 - Depth, each layer $h^{(i)}$ has d elements, each of which identifies 2 region inputs, so each layer has 2^d regions that are identified.

Since the entire net has L layers, and from the result of 3.2, composition of functions identifies a number of regions equal to the product of the number of regions identified by each composed function,

Then the L layer net is a composition of L functions and each layer has $\underbrace{2 \cdot 2 \cdots \cdot 2}_{d}$ or 2^d identified regions

$$L \text{ layer net} = \underbrace{2^d \cdot 2^d \cdot \ldots \cdot 2^d}_{L} = 2^{Ld}$$