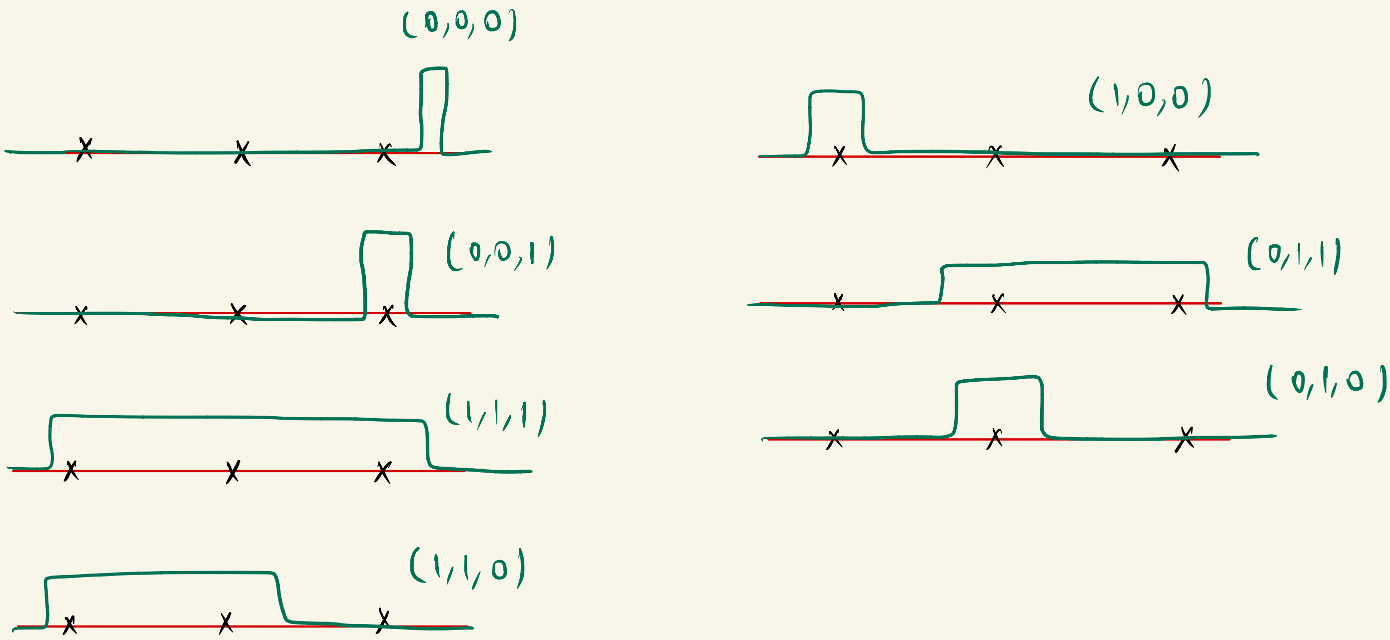
Problem 1

(a-i) $\quad C_1 = \{1,2\}$

$$\mathcal{H}_{C_1} = \{(0,0),\ (0,1),\ (1,0),\ (1,1)\}$$

(a-ii) $\quad C_2 = \{1,2,3\} \qquad \mathcal{H}_{C_2} = \{(0,0,0),\ (0,0,1),\ (0,1,1),$
$(1,1,1),\ (1,1,0),\ (1,0,0),\ (0,1,0)\}$



Note that we cannot generate $(1,0,1)$ using any function in $\mathcal{H}$.

(b) $C = \{c_1, c_2, c_3\}$ where $c_1 < c_2 < c_3$. We cannot generate the label $(1,0,1) \rightarrow |\mathcal{H}_C| < 8$.

(c) $C_1 = \{1,2\}$ is shattered by $\mathcal{H}$.

$C_2 = \{1,2,3\}$ is not shattered by $\mathcal{H}$. $\Rightarrow VC(\mathcal{H}) \geq 2$

We showed that no set of size 3 can be shattered. $\Rightarrow VC(\mathcal{H}) < 3$

$\Rightarrow \boxed{VC(\mathcal{H}) = 2}$ ∎

<u>Problem 2</u> Let C be a set that can be shattered using $\mathcal{H}_1$ and $|C| = VC(\mathcal{H}_1)$.

$\Rightarrow$ For any labeling on the members of C, there exists $h \in \mathcal{H}_1$ that generates that labeling.

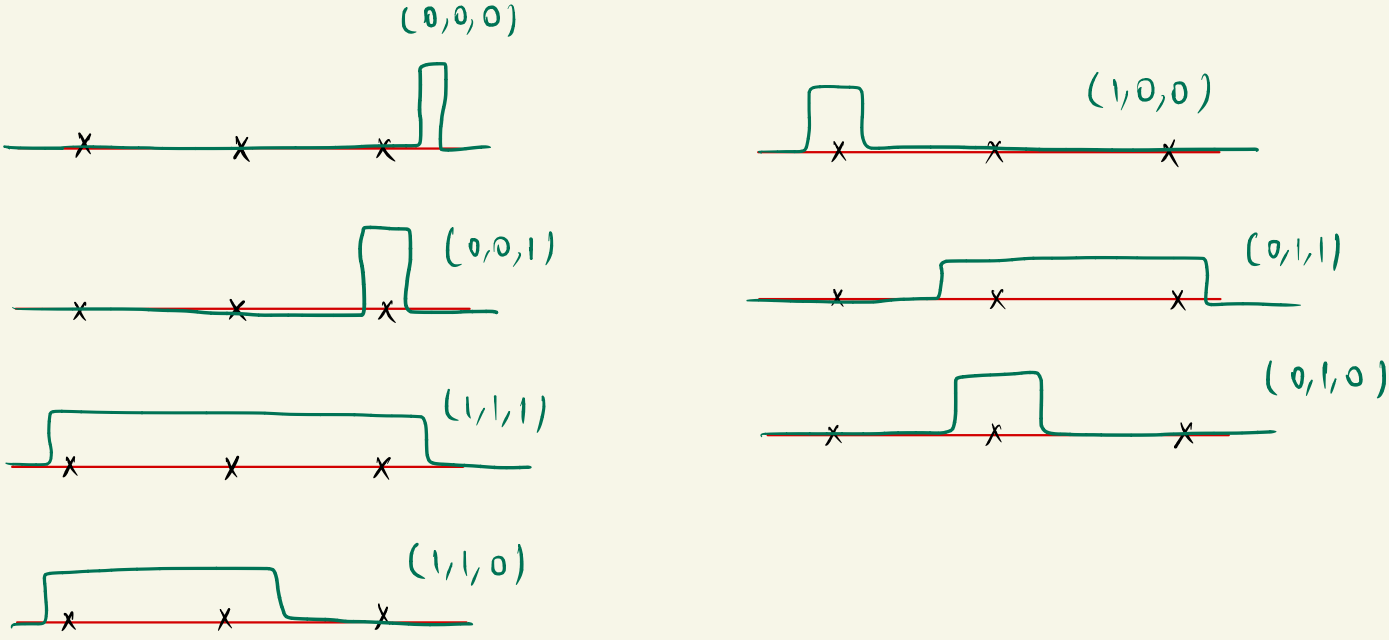
$h \in \mathcal{H}_1 \subseteq \mathcal{H}_2 \Rightarrow h \in \mathcal{H}_2 \Rightarrow \mathcal{H}_2$ also shatters C.

$\Rightarrow VC(\mathcal{H}_2) \geq VC(\mathcal{H}_1)$.

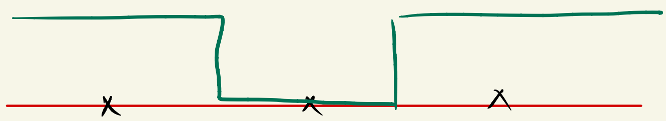
<u>Problem 4</u> $\mathcal{H}$ = class of signed intervals.

Claim: $VC(\mathcal{H}) = 3$.

<u>Part a</u>: We show that it can shatter a set of size 3. Let $C = \{c_1, c_2, c_3\}$ $c_1 < c_2 < c_3$.



However, H can also generate (1,0,1):

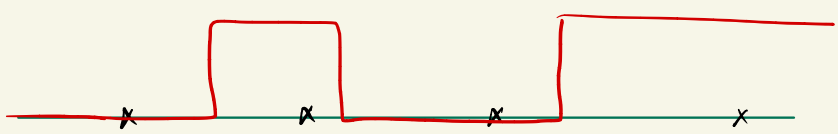


$\Rightarrow |H_C| = 8 = 2^{|C|} \Rightarrow C$ is shattered

$\Rightarrow VC(H) \geq 3.$

Part b We then show that it cannot shatter any set of size 4.

Let $C = \{c_1, c_2, c_3, c_4\}$. H cannot generate the labeling (0,1,0,1)
$c_1 < c_2 < c_3 < c_4$



$\rightsquigarrow$ not possible

$\Rightarrow VC(H) < 4.$

Part a + part b $\Rightarrow$ $VC(\mathcal{H}) = 3.$

Problem 5

(a): Look at the lecture note.

(b): Any set can be shattered because any labeling can be generated $\Rightarrow$ $VC = \infty$.

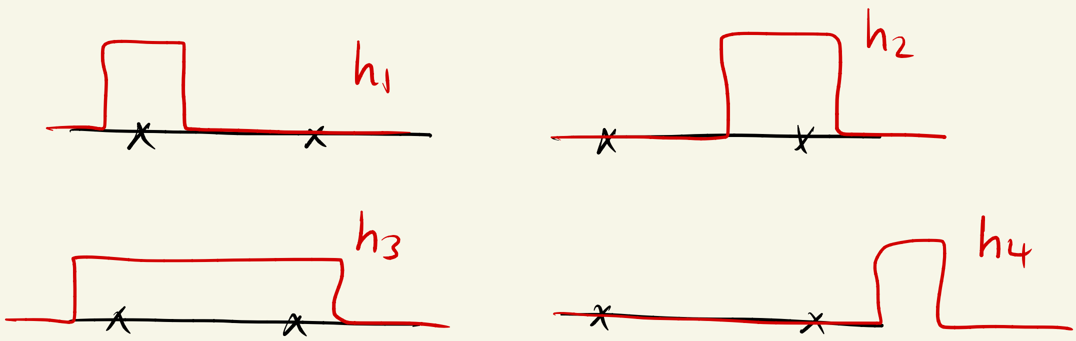
(c): Let C such that $|C| > \log m$.

$2^{|C|} > 2^{\log|\mathcal{H}|} = |\mathcal{H}|$ $\Rightarrow$ there are not enough functions to generate all $2^{|C|}$ possible distinct labelings.

$\Rightarrow$ $VC < \log|\mathcal{H}|$.

There are hypothesis classes that achieve this

For example let $|\mathcal{H}| = 4 = 2^2$



$\mathcal{H} = \{h_1, h_2, h_3, h_4\}$

(d) To show $VC = d$, we show

(1) there exists a set of size d that is shattered.

(2) show there is no set of size d+1 that are shattered.

Problem 3 (20 points) Consider the following hypothesis class

$$\mathcal{H} = \{h_{w,b} \mid w, b \in \mathbb{R}\},$$

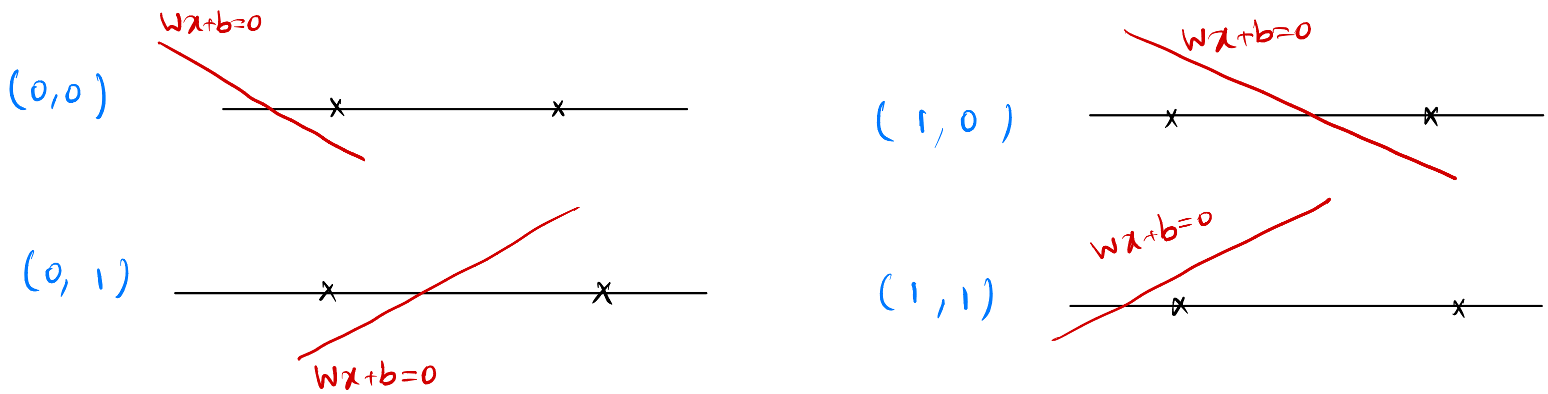
in which

$$h_{w,b}(x) = \begin{cases} 1 & wx + b \geq 0 \\ 0 & wx + b < 0 \end{cases}$$

In this problem, we want to show that the VC dimension of $\mathcal{H}$ is 2.

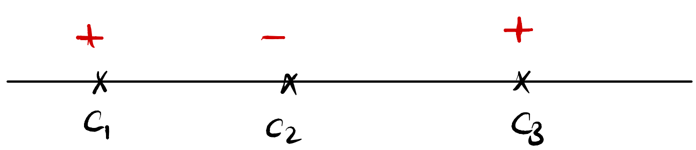
1. [10pts] Show that there exists a set of size 2 that can be shattered by $\mathcal{H}$.

Consider $C = \{-1, +1\}$. This set can be shattered.



2. [10pts] Show that no set of size 3 can be shattered by $\mathcal{H}$.

Consider <u>any</u> set $C = \{c_1, c_2, c_3\}$. WLOG assume $c_1 < c_2 < c_3$:



We cannot create $+ - +$

because $wx+b$ is monotone. (or a similar argument)