Exercise

Let

$$\mathcal{H}_d = \{h_{\mathbf{w}}(\mathbf{x}) : h_{\mathbf{w}}(\mathbf{x}) = \text{sign}(< \mathbf{w}, \mathbf{x} >)\}$$

where $\mathcal{X} = \mathbb{R}^d$.

Prove that $VCdim(\mathcal{H}_d) = d$.

Solution We need to prove that $VCdim(\mathcal{H}_d) \geq d$ and that $VCdim(\mathcal{H}_d) \leq d$.

i) $VCdim(\mathcal{H}_d) \geq d$. We need to show a set of d vectors in $\mathbb{R}^d$ that is shattered by $\mathcal{H}_d$.

Consider $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_d\}$ with $\vec{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} i$, $\forall 1 \leq i \leq d$

This set is shattered by $\mathcal{H}_d$: we need to show that for every labeling $y_1, y_2, \dots, y_d$, where y_i is the label of $\vec{e}_i$, with $y_i \in \{-1, 1\}$, there is an hypothesis in $\mathcal{H}_d$ that

assigns such labels to the set.
Consider an arbitrary labeling $y_1, y_2, \dots, y_d$: consider the hypothesis $h_{\vec{w}}$ where $\vec{w} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{bmatrix}$. We have that for every i, with $1 \leq i \leq d$:

$$h_{\vec{w}}(\vec{e}_i) = \text{sign}(\langle \vec{w}, \vec{e}_i \rangle) = \text{sign}\left(\left\langle \begin{bmatrix} y_1 \\ \vdots \\ y_d \end{bmatrix}, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\rangle \right) = \text{sign}(y_i) = y_i$$

ii) $VCdim(\mathcal{H}_d) \leq d$: we need to show that no set of $d+1$ vectors in $\mathbb{R}^d$ can be shattered by $\mathcal{H}_d$.
Consider an arbitrary set $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{d+1}\}$ with $\vec{x}_i \in \mathbb{R}^d$ for $1 \leq i \leq d+1$.
They cannot be linearly independent $\Rightarrow \exists\, a_1, a_2, \dots, a_{d+1}$ with $a_i \in \mathbb{R}$, $1 \leq i \leq d+1$, such that:

- not all a_i's are 0 $(\star)$

$\Rightarrow \sum_{i=1}^{d+1} a_i \vec{x}_i = 0$ $(\star\star)$

Define: $I = \{ i : a_i > 0 \}$, $J = \{ j : a_j < 0 \}$. Note that it cannot be that $I = \emptyset = J$ (due to $(\star)$)

There are 3 cases: i) $I \neq \emptyset \neq J$; ii) $I \neq \emptyset = J$; iii) $I = \emptyset \neq J$

Case i) we are assuming $I \neq \emptyset \neq J$. Then

$$\underbrace{\sum_{i \in I} a_i \vec{x}_i = \sum_{j \in J} |a_j| \vec{x}_j}_{(\star\star\star)}$$

$$\sum_{i=1}^{d+1} a_i \vec{x}_i = \sum_{i \in I} a_i \vec{x}_i + \sum_{j \in J} a_j \vec{x}_j = 0 \quad (\text{by } \star\star)$$

$$\Longleftrightarrow \sum_{i \in I} a_i \vec{x}_i = \sum_{j \in J} - a_j \vec{x}_j$$

$$\Longleftrightarrow \sum_{i \in I} a_i \vec{x}_i = \sum_{j \in J} |a_j| \vec{x}_j$$

Assume that $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{d+1}\}$ is shattered by $\mathcal{H}_d$:
must exists $\vec{w}$ such that

$$\langle \vec{w}, \vec{x}_i \rangle > 0 \quad \forall\, i \in I$$
$$\langle \vec{w}, \vec{x}_j \rangle < 0 \quad \forall\, j \in J$$

$$\begin{aligned} 0 < \sum_{i \in I} \underbrace{a_i}_{>0} \underbrace{\langle \vec{w}, \vec{x}_i \rangle}_{>0} &= \langle \overbrace{\sum_{i \in I} a_i \vec{x}_i}^{(**)}, \vec{w} \rangle \\ &= \langle \sum_{j \in J} |a_j| \vec{x}_j, \vec{w} \rangle \\ &= \sum_{j \in J} \underbrace{|a_j|}_{>0} \underbrace{\langle \vec{x}_j, \vec{w} \rangle}_{<0} \\ &< 0 \end{aligned}$$

$\Rightarrow$ contradiction

Case ii) : $I \neq \emptyset = J$: same steps lead to

$$0 < \ldots \leq 0 \Rightarrow \text{contradiction}$$

Case iii) : $I = \emptyset \neq J$: same steps lead to

$$0 \leq \ldots < 0 \Rightarrow \text{contradiction}$$

$\square$