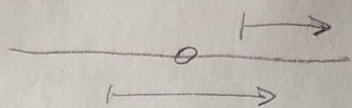
ECE 595 HW6

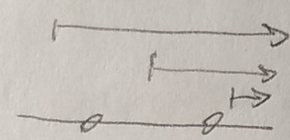
1. a) $H = \{h : \mathbb{R} \to \{-1,+1\} \mid h(x) = +1,\ \forall x \in [a, \infty),\ a \in \mathbb{R}\} \sqcup$
$\{h : \mathbb{R} \to \{-1,+1\} \mid h(x) = +1,\ \forall x \in (-\infty, a],\ a \in \mathbb{R}\}$

$N=1$



$m_H(1) = 2$

$N=2$

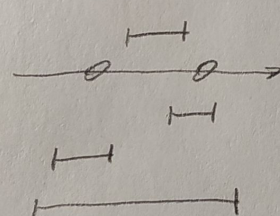


$m_H(2) = 3$

$d_{vc} = 1$

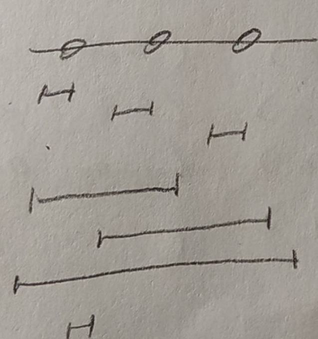
b) $H = \{h : \mathbb{R} \to \{-1,1\} \mid h(x) = 1,\ \forall x \in [a,b],\ a,b \in \mathbb{R}\} \sqcup$
$\{h : \mathbb{R} \to \{-1,1\} \mid h(x) = -1,\ \forall x \in [a,b],\ a,b \in \mathbb{R}\}$

$N=2$



$m_H(2) = 4$

$N=3$

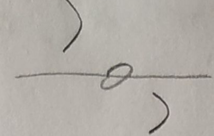


$m_H(3) = 7$

$\therefore d_{vc} = 2$

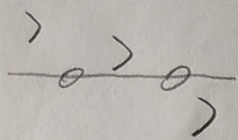
c) $H = \{h: \mathbb{R}^2 \to \{-1,+1\} \mid h(x) = +1,\ \forall x \text{ where } \|x\|_2 \leq b,\ b \in \mathbb{R}\}$

$N=1$



$m_H(1) = 2$

$N=2$

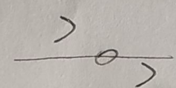


$m_H(2) = 3$

$\therefore\ d_{vc} = 1$

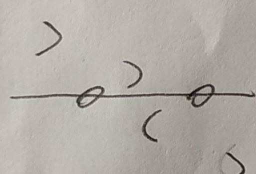
d) $H = \{h: \mathbb{R}^2 \to \{-1,+1\} \mid h(x) = +1,\ \forall x \text{ where } \|x-a\|_2 \leq b,$
$a \in \mathbb{R}^2,\ b \in \mathbb{R}\}$

$N=1$



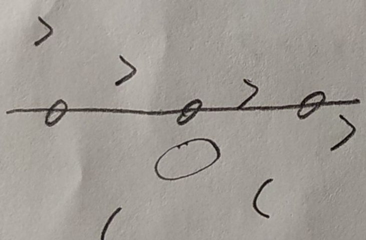
$m_H(1) = 2$

$N=2$



$m_H(2) = 4$

$N=3$



(0, 0, 0)	(0, 0, 1)
(1, 0, 0)	(0, 1, 1)
(1, 1, 0)	~~(1, 0~~ 0, 1, 0)
(1, 1, 1)	(

$m_H(3) = 7$

$d_{vc} = 2$

2 $\alpha \in \mathbb{R}$

$H = \{ h_\alpha : \mathbb{R} \rightarrow \mathbb{R} \mid h_\alpha(x) = (-1)^{\lfloor \alpha x \rfloor}, \alpha \in \mathbb{R} \}$

$(x_1, x_2, \cdots x_N) = (10^0, 10^1, \cdots, 10^{N-1})$

Since $\alpha \in \mathbb{R}$

$\lfloor \alpha x_i \rfloor$ can be either odd or even.

$\therefore$ $m_H(N) = 2^N$

$d_{vc} = \infty$.

3 $y_n = x_n^T \theta + e_n$ $\quad n = 1, \cdots, N$ $\quad e_n \sim$ Gaussian $(0, \sigma^2)$

$y = X \underline{\theta} + \underline{e}$ $\longrightarrow y_n = \underline{x}_n^T \underline{\theta} + e_n$

$D = \{(x_1, y_1), \cdots (x_N, y_N)\}$

a) $g^{(D)}(x') = \hat{\theta}^T x'$ $\quad$ L_2-loss

$g^{(D)}(x') = \cancel{\theta X - e}\ X\theta - e$ $\quad$ $\cancel{g_\theta(x) = \theta^T x}$

$\hat{\theta} = \underset{\underline{\theta}}{\arg\min} \| \underline{y} - X \underline{\theta} \|^2$ $\quad$ $g_{\underline{\theta}}(\underline{x}) = \underline{\theta}^T \underline{x}$

$= (X^T X)^{-1} X^T \underline{y}$

$g^{(D)}(\underline{x}')$

$= \hat{\underline{\theta}}^T \underline{x}' = (\underline{x}')^T \hat{\underline{\theta}} = (\underline{x}')^T [(X^T X)^{-1} X^T \underline{y}]$

$= (\underline{x}')^T [(X^T X)^{-1} X^T (X\underline{\theta} + \underline{e})]$

$= (\underline{x}')^T [\underline{\theta} + (X^T X)^{-1} X^T \underline{e}] = (\underline{x}')^T \underline{\theta} + (\underline{x}')^T (X^T X)^{-1} X^T \underline{e}$

b) $\bar{g}(x') \overset{def}{=} \mathbb{E}_D [g^{(D)}(x')]$

$= \mathbb{E}_e [(\underline{x}')^T \underline{\theta} + (\underline{x}')^T (X^T X)^{-1} X^T \underline{e}]$

$= (\underline{x}')^T \underline{\theta} + (\underline{x}')^T (X^T X)^{-1} X^T \cancel{\mathbb{E}_{\underline{e}}[\underline{e}]} \to 0$

$= (\underline{x}')^T \underline{\theta}$

Yes. Same form of the original model

c) $E_D[(g^{(D)}(x') - \bar{g}(x'))^2] = \cancel{\sigma^2 (1 - \frac{d}{N})}$

$= E_e[(\cancel{(\underline{x}')^T \underline{\theta}} + (\underline{x}')^T (X^T X)^{-1} X^T \underline{e} - \cancel{(\underline{x}')^T \underline{\theta}})^2]$

$= E_e[((\underline{x}')^T (X^T X)^{-1} X^T \underline{e})^2] = ((\underline{x}')^T (X^T X)^{-1} X^T)^2 E_e[(\underline{e})^2]$

$= ((\underline{x}')^T (X^T X)^{-1} X^T)^2 \sigma^2$