PAC Learning

$$H = \{ h_1 , \dots , h_M \}$$

Output hypothesis = $g \in H$

$$\Pr (|E_{in}(g) - E_{out}(g)| > \varepsilon) \leq 2Me^{-2N\varepsilon^2}$$

$$E_{out}(g) \leq E_{in}(g) + \sqrt{\frac{1}{2N} \log \frac{2M}{\delta}}$$

$\longrightarrow$ Holds with probability $\geq 1 - \delta$

Chapter 2, $M = \infty$

Main step in previous analysis

$$\Pr (|E_{in}(g) - E_{out}(g)| > \varepsilon) \leq$$

$$\Pr (\bigcup_{h \in H} |E_{in}(g) - E_{out}(g)| > \varepsilon) \leq$$

$$\sum_{h \in H} \Pr (|E_{in}(g) - E_{out}(g)| > \varepsilon)$$ Every Loose

Dichotomy Vector

$h \in H \qquad \underline{x_1 , \dots , x_N} \in \mathbb{R}^d$

$(h(\underline{x_1}), \dots , h(\underline{x_N})) \in \{-1, +1\}^N$

$h : \mathbb{R}^d \longrightarrow \{ -1 , +1 \}$

Is known as Dichotomy Vector of h

Recall

In Sample Training Error

$$E_{in}(g) = \frac{1}{N}\sum_{n=1}^{N} \underline{1}\{y_n \neq g(\underline{x_n})\}$$

Test Error

$$E_{out}(g) = \Pr(y \neq g(\underline{x}))$$

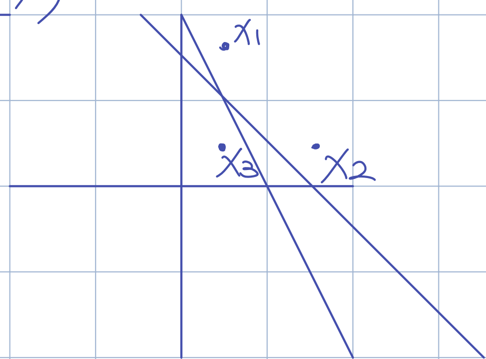
$E_{in}(h)$ depends on $h(\cdot)$ through
$\underline{(h(x_1), \ldots, h(x_N))}$

If $h_1(\cdot) \neq h_2(\cdot)$

but $h_1(\underline{x_1}) = h_2(\underline{x_1})$

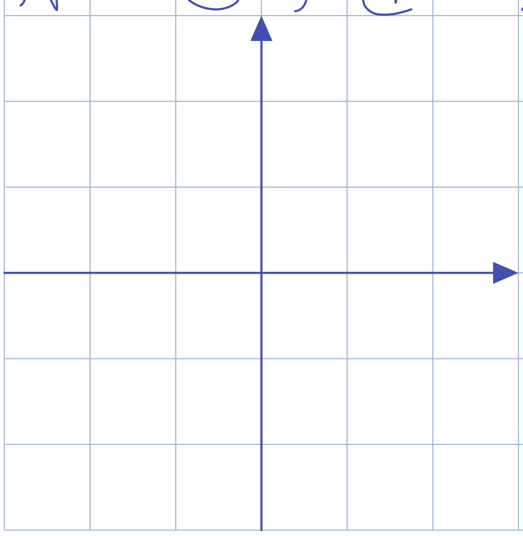
$\vdots$

$h_1(\underline{x_N}) = h_2(\underline{x_N})$



Example

$N = 3, \; d = 2$



$$H = \{h_1, \ldots, h_N\}$$

$$\{h_1(\underline{x_1}), \ldots, h_1(\underline{x_3})\}$$

$$\vdots$$

$$\{h_N(\underline{x_1}), \ldots, h_N(\underline{x_3})\}$$

Goal

$$\sum_{\substack{h \\ E_{in}(h)\ \text{distinct}}} \Pr\left(\left| E_{in}(h) - E_{out}(h) \right| > \varepsilon \right) -$$

$$\sum_{\substack{h: \\ \text{Distinct Dichotomy Vector}}} \Pr\left(\left| E_{in}(h) - E_{out}(h) \right| > \varepsilon \right)$$

# of distinct dichotomy vectors is $2^3 = 8$

Dichotomy Set

$$H(x_1, \ldots, x_N) = \{ (h(\underline{x_1}), h(\underline{x_2}), \ldots, h(\underline{x_N})) : h \in H \}$$

$$|H(\underline{x_1, \ldots, x_N})| \leq 2^N$$

Definition: Hypothesis Class H Shatters $\underline{x_1, \ldots, x_N}$ if $|H(\underline{x_1, \ldots, x_N})| = 2^N$

That is H can attain all dichotomy vector on $\underline{x_1, \ldots, x_N}$

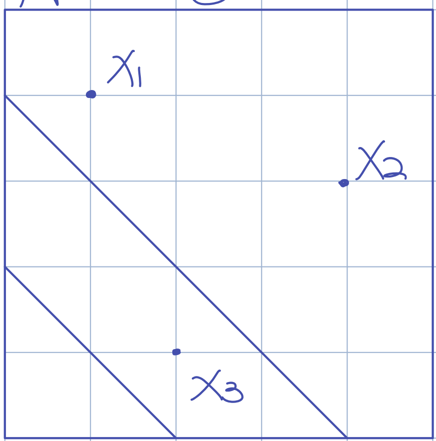
Example,

$H \equiv$ set of all linear classifier in 2D

$\hat{y} = \text{sign}(w_0 + w_1 x_1 + w_2 x_2)$

$w_0, w_1, w_2 \in \mathbb{R}$

$N = 3$



(+++)

+	+	+
+	+	−
+	−	+
+	−	−

−	+	+
−	+	−
−	−	+
−	−	−

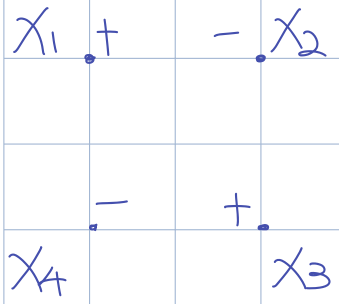
$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & x_1(1) & x_2(1) \\ 1 & x_1(2) & x_2(2) \\ 1 & x_1(3) & x_2(3) \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix}$$

Conclusion: H will shatter x_1, x_2, x_3 unless they lie on a straight line so that the matrix

$$\begin{bmatrix} 1 & x_1(1) & x_2(1) \\ 1 & x_1(2) & x_2(2) \\ 1 & x_1(3) & x_2(3) \end{bmatrix}$$ Is Not Invertible

$N = 4$, $d = 2$

$H \equiv$ Linear Classifier



$\hat{y} = \text{sign}(w_0 + w_1 x_1 + w_2 x_2)$

$|H(\underline{x_1, \dots, x_N})| = 14$

H does not shatters $\underline{x_1, x_2, x_3, x_4}$