Minimax lower bounds (continued)
Chapter 11

Recall

$H_0 \quad P_0 \sim Y$ $\qquad\qquad \rho = \sum_y \sqrt{P_0(y) P_1(y)}$

vs $H_1 \quad P_1 \sim Y$

Lemma $\qquad \dfrac{\rho^2}{4} \le P_e^* \le \dfrac{\rho}{2}$

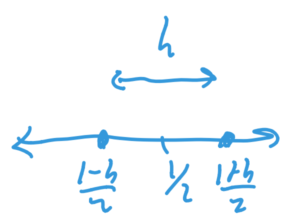
$\uparrow$ minimum average probability of error

Example

$Y = (Y_1, \dots, Y_n) \sim Be\left(\frac{1-h}{2}\right) \quad H_0 \qquad 1 \ge h \ge 0$

$\phantom{Y = (Y_1, \dots, Y_n) \sim{}} Be\left(\frac{1+h}{2}\right) \quad H_1$

$0 \quad 1$



$n=1 \quad P_0 = \left(\frac{1+h}{2}, \frac{1-h}{2}\right)$

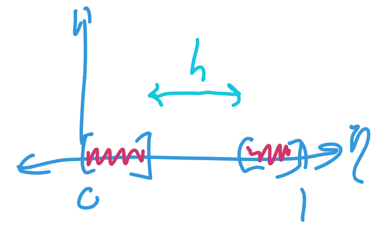
$ P_1 = \left(\frac{1-h}{2}, \frac{1+h}{2}\right) \qquad \rho = 2\sqrt{\frac{1-h^2}{4}} = \sqrt{1-h^2}$

$n \ge 1 \qquad \rho = \left(\sqrt{1-h^2}\right)^n \to 0$

$\rho^2 = (1-h^2)^n \qquad\qquad \dfrac{(1-h^2)^n}{4} \le P_e^* \le \dfrac{(1-h^2)^{n/2}}{2}$

$\left(h = \frac{1}{\sqrt{n}} \quad (1-h^2)^n = \left(1 - \frac{1}{n}\right)^n \approx e^{-1} \right.$

$\mathcal{P}(h)$ — P such that for any x,
if $\eta(x) = P\{Y=1 \mid X=x\}$, then
$$\eta(x) \in \left[0, \tfrac{1-h}{2}\right] \cup \left[\tfrac{1+h}{2}, 1\right]$$



Thm 12.1 Let $\mathcal{F}$ have VC dimension $V \geq 2$. Then for any $n \geq V$, and $h \in [0,1]$

$$R_n(h, \mathcal{F}) \geq c \min\left(\sqrt{\frac{V}{n}}, \frac{V}{nh}\right) \quad \left(c = \tfrac{1}{32} \text{ works}\right)$$

↑ min max expected excess risk

12.2 $\mathcal{X}$, $\mathcal{F}: \mathcal{X} \to \{0,1\}$, $P \in \mathcal{P}(h)$

$$L(f) = P\{Y \neq f(X)\}$$

$$f^*(x) = 1_{\{\eta(x) \geq 1/2\}} \quad \text{(MAP estimator)}$$

$$R_n(h, \mathcal{F}) \triangleq \inf_{A} \sup_{P \in \mathcal{P}(h)} E\left[L(\hat{f}_n) - L^*\right] \quad \left(\text{min max}\right)$$

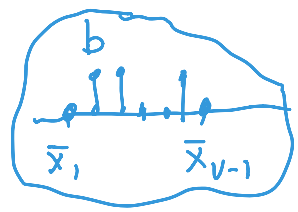
$$\hat{f}_n = A(Z^n)$$
$$Z^n = ((X_1, Y_1), \ldots (X_n, Y_n))$$

Proof Let $\bar{x}_1, \ldots, \bar{x}_V$ be V points in $\mathcal{X}$ that are shattered by $\mathcal{F}$.

For any $b \in \{0,1\}^{V-1}$ there exists $f_b \in \mathcal{F}$

such that $f_b(\bar{x}_v) = b_v \quad 1 \le v \le V-1$

(Let $b_V = 0$) $\quad f_b(\bar{x}_V) = 0$



$\frac{1-h}{2} \quad \frac{1+h}{2}$

Use flip probability $\frac{1-h}{2}$.

Let P_b denote the joint dist$^{\underline{n}}$ on $X \times \{0,1\}$ so

if (X,Y) has dist$^{\underline{n}}$ P_b then

$$X = \begin{cases} \bar{x}_v & \text{w/p } \frac{p}{V-1} \text{ if } 1 \le v \le V-1 \\ = \bar{x}_V & \quad " \quad 1-p \end{cases}$$

b_v

Given $X = \bar{x}_v$, $P_b(Y=1 \mid X=\bar{x}_v) = \underline{\underline{\frac{1-h}{2} + h\,\mathbb{1}_{\{b_v=1\}}}}$

Defines 2^{V-1} probability distributions on $X \times \{0,1\}$.

Let $\bar{P}^n$ be the joint dist$^{\underline{n}}$ of (B, Z^n) as follows.

where $Z^n = ((X_1,Y_1), \cdots, (X_n,Y_n))$

Let B be uniformly distributed over $\{0,1\}^{V-1}$

Given B, let $\underline{Z^n \mid B} = b$ have dist$^{\underline{n}}$ P_b.

Let $\hat{f}_n$ be output of an arbitrary algorithm.

$$L_B(\hat{f}_n) = \frac{1-h}{2} + \frac{ph}{V-1} \sum_{v=1}^{V-1} \mathbb{1}\{\hat{f}_n(\bar{x}_v) \neq B_v\}$$

$$\mathbb{E}^n\left[L_B(\hat{f}_n)\right] - \frac{1-h}{2} = \frac{ph}{V-1} \sum_{v=1}^{V-1} P\{\hat{f}_n(\bar{x}_v) \neq B_v\}$$

expected excess risk if P_b is true dist^{n} per sample

$$\frac{1}{2^{V-1}} \sum_{b \in \{0,1\}^{V-1}} \left[\mathbb{E}_b^n\left[L_b(\hat{f}_n)\right] - \underbrace{\frac{1-h}{2}}_{L^*}\right] = \frac{ph}{V-1} \sum_{v=1}^{V-1} P\{\hat{f}_n(\bar{x}_v) \neq b_v\}$$

$$\therefore R_n(h, \mathcal{F}) \geq \sup_b \qquad \geq ph\left(\frac{1}{V-1}\right) \sum_{v=1}^{V-1} \underbrace{P\{\hat{f}_n(\bar{x}_v) \neq b_v\}}_{\text{Lower bound.}}$$

Let N_v = # of X_i's equal to $\bar{x}_v$. $1 \leq v \leq V-1$

By Bhattacharyya bound

$$\bar{P}^n\{\hat{f}_n(\bar{x}_v) \neq B_v \mid N_v\} \geq \frac{1}{4}(1-h^2)^{N_v}$$

$$\mathbb{E}[N_v] \overset{?}{=} \frac{np}{V-1} \qquad N_v \overset{d}{=} U_1 + \dots + U_n \quad U_i \sim \mathrm{Ber}\left(\frac{p}{V-1}\right)$$

So $$\bar{P}^n\{\hat{f}_n(\bar{x}_n) \neq B_v\} = \mathbb{E}\left[\qquad\right] =$$

$$= \frac{1}{4}\left(\frac{p}{V-1}(1-h^2) + \frac{1-p}{V-1}\right)^n$$

$$= \frac{1}{4}\left(1 - \frac{ph^2}{V-1}\right)^n$$

$$\boxed{\begin{array}{l}\text{So} \\ (*)\ R_n(h,f) \geq \frac{ph}{4}\left(1-\frac{ph^2}{V-1}\right)^n\end{array}}$$ all $n \geq 1$, $V \geq 2$, $h \in [0,1]$, $p \in [0,1]$

<u>Case 1</u> $\sqrt{\frac{V-1}{n}} \leq h \leq 1$

Let $p = \frac{V-1}{(n+1)h^2}$. Then

$(*)$ yields $R_n(h,f) \geq \frac{V-1}{4(n+1)h}\left(1-\frac{1}{n+1}\right)^n \geq \frac{V-1}{16hn}$

$$\geq \frac{V}{32hn}$$

<u>Case 2</u> $0 \leq h \leq \sqrt{\frac{V-1}{n}}$ — use bound for case 1 with $h = \sqrt{\frac{V-1}{n}}$.