CSCI 5622
Homework 4 – Learnability
Nick Ketz

- Sorry I didn't have time to write the solutions up in Latex, hopefully everything is legible.

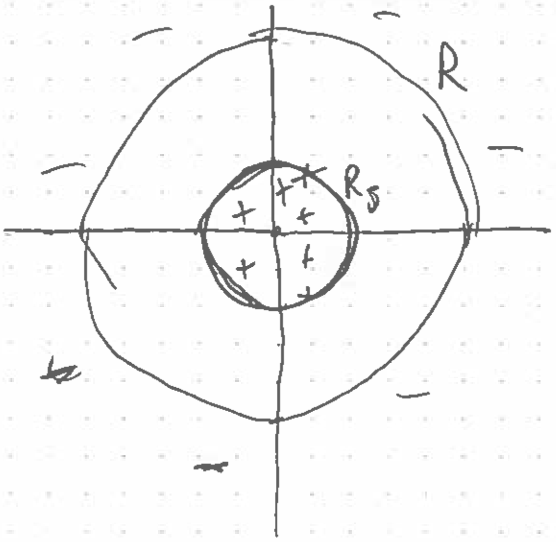
- My math notation is loose at best, so I usually tried to express in words what my sometimes-ambiguous notation was attempting to convey.

- Problem 3.19(e) was left incomplete as I was unable to come up with a good argument in the allotted time

- Problem 3.12(b) is also incomplete. I knew what I needed to show, but was unable to derive a proof.

2.3

$C = \{(x,y) : x^2 + y^2 \le r^2\} \Rightarrow m \ge 1/\varepsilon \log 1/\delta$



$R_s \Rightarrow$ smallest circle defined by m points

$R \Rightarrow$ target concept, $\Pr[\mathcal{R}(R) \le \varepsilon] = 1$

$\Pr[R_s] \Rightarrow$ probability mass of region defined by R_s

$\mathcal{R}(R_s) =$ generalization error of R_s

$$\Pr_{S\sim D^m}[\mathcal{R}(R_s) > \varepsilon] \le \Pr_{S\sim D^m}[\{R_s \cap R = \emptyset\}]$$
$$\le (1-\varepsilon)^m$$
$$\le e^{\wedge}[-m\varepsilon]$$

if generalization error is less than the mass of R that is not contained in R_s, which is less than the mass of space outside R i.e. $(1-\varepsilon)$

$$\Pr_{S\sim D^m}[\mathcal{R}(R_s) > \varepsilon] \le \delta$$
$$e^{\wedge}[-m\varepsilon] \le \delta$$

$$-m\varepsilon \le \log \delta$$
$$m \ge -1/\varepsilon \log \delta$$
$$m \ge -1/\varepsilon \log 1/\delta$$

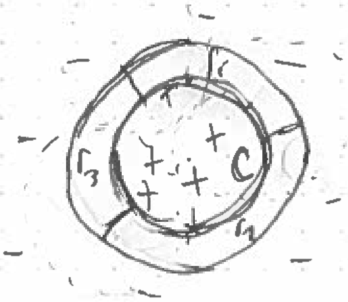
2.4

$C = \{ x \in \mathbb{R}^2 : \| x - x_0 \| \leq r \}$

two free parameters $x_0 \in \mathbb{R}^2$ and $r \in \mathbb{R}$

Propose $m \geq (3/\epsilon) \log(3/\delta) \Rightarrow \Pr[R(C_s) > \epsilon]$

$\leq \Pr[\cup_i \{ C_s \cap r_i = \emptyset \}]$



$\Pr[r_i] \geq \epsilon/3$

C_s = concept based on current sample S

C_T = target concept

Shaded region $\Rightarrow$ mass that contributes to $\Pr[R(C_s) > \epsilon]$

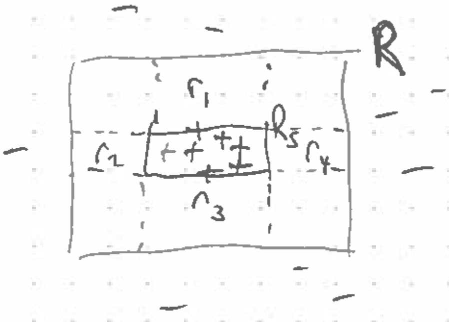
- Solution does not consider x_0 variation, i.e. target concept C_T may not be centered on the same point as C_s and error can also be due to false positives (negative points falling in missaligned C_s)

therefore

$\Pr[R(C_s) > \epsilon] \geq \Pr[\cup_i \{ r_i \cap C_s = \emptyset \}]$

mass of geralization error is not only attributable to the mass of r_i regions

2.6 noise $\eta \in (0, \frac{1}{2})$ $\eta \leq \eta' < \frac{1}{2}$



Non noisy case

$Pr[R] > \varepsilon$

if $R_m(R') > \varepsilon \Rightarrow \sum_{i=1}^{4} Pr[\{R' \cap r_i = \emptyset\}]$
$\leq 4(1 - \varepsilon/4)^m$

With noise η': negative points unaffected, positive points flip with probability η'

for noisy case need to consider errors due to flipped labels outside ε in addition to true points that fall in ε.

a) $\sum_i^4 \left(Pr[\{R' \cap r_i = \emptyset\}] \text{ or } Pr[\text{flipped label inside } r_i] \right)$

$\leq \sum_i^4 \left[\left((1 - \varepsilon/4) + \varepsilon/4 \, \eta' \right)^m \right]$

$\leq 4(1 - \varepsilon/4(1-\eta'))^m = 4 \exp[-m \, \varepsilon/4(1-\eta')]$

b) $\delta \geq 4 \exp[-m \, \varepsilon/4 (1-\eta')]$

$\log \delta/4 \geq -m \, \varepsilon/4 (1-\eta')$

$\frac{4}{\varepsilon(1-\eta')} \log 4/\delta \leq m$

3.5 Propose $\mathcal{R}_m(H) \leq \mathcal{O}\left(\frac{VCDim(H)}{m}\right)$

Suppose $|H_2| = 2 \Rightarrow$ all $+1$, and all -1

$VCDim(H_2) = 1 \Rightarrow \mathcal{R}_m(H_2) \leq \mathcal{O}\left(\frac{1}{m}\right)$

From Massart's Lemma and 3.23 in textbook

$$\mathcal{R}_m(H) \leq \sqrt{\frac{2\log(\Pi_H(m))}{m}}$$

From Sauer's lemma and 3.3 in textbook

$$\Pi_H(m) \leq \mathcal{O}(m^d), \quad d < \infty$$

$\mathcal{R}_m(H_2) \leq \mathcal{O}\left(\frac{1}{m}\right)$ Proposal

$$\mathcal{R}_m(H_2) \leq \mathcal{O}\sqrt{\frac{2\log(\Pi_m(H_2))}{m}} \text{ from 3.23}$$

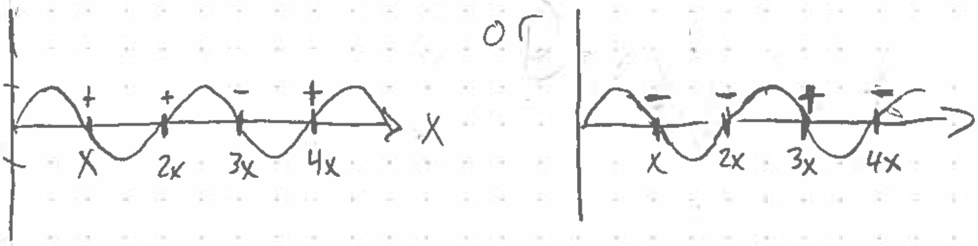
$$\leq \mathcal{O}\sqrt{\frac{2\log \mathcal{O}(m^1)}{m}} \text{ from 3.3}$$

$$\simeq \mathcal{O}\left(\sqrt{\frac{\log m}{m}}\right)$$

Contridiction: $\mathcal{O}\left(\sqrt{\frac{\log m}{m}}\right) \geq \mathcal{O}\left(\frac{VCDim(H_2)}{m}\right)$

3.12 $\{x \to \sin(\omega x) : \omega \in \mathbb{R}\}$.

a) show $x \in \mathbb{R}$ $x, 2x, 3x, 4x$ is not shatterable



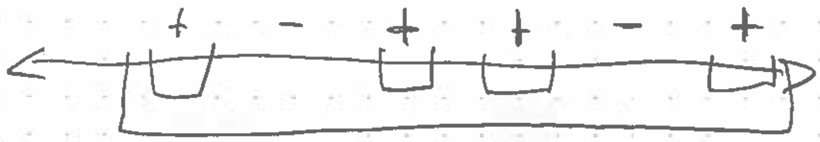
because sin function is periodic there can always be a set of points that are placed equi-distance from each other with non-periodic labels that can not be shattered, given a consistent rule for intersection with the sin curve

b) if points are distributed as 2^{-m}, sin function with sufficiently high ω can correctly classify any labeling

$\frac{1}{2}(2\pi\omega)$ = min length of contiguous sequence of a single class

$\omega = \frac{1}{\pi} \min[I(Y)]$, I returns length of contiguous sequences in Y

3.13 VC Dim of subsets of real line formed by union of K-intervals



any labeling of n positive points can be given an interval of its own, and then an interval spanning all intervals will create a union across them, shattering any set with n positive labels using $n+1$ intervals

$n \to \infty \Rightarrow$ VC Dim $= \infty$

3.19

$Pr[X_A = 0] = 1/2 - \varepsilon/2 = p_A \quad q_A = 1/2 + \varepsilon/2$

$Pr[X_B = 0] = 1/2 + \varepsilon/2 = p_B \quad q_B = 1/2 - \varepsilon/2$

$X \in \{X_A, X_B\}$ uniform random draw X_i. then $S: \{0,1\}^m$ from X_i

a) $f_0 : S \to \{X_A, X_B\}$

$f_0(S) = X_A$ iff $N(S) < m/2$

Show: $\text{error}(f_0) = R(f_0) \geq 1/2\, Pr[N(S) \geq m/2 \mid X = X_A]$

if m = even,
- errors happen when:

$X = X_A$ & $N(S) \geq m/2 = Pr[N(S) \geq m/2 \mid X = X_A]$

or $X = X_B$ & $N(S) < m/2 = Pr[N(S) < m/2 \mid X = X_B]$

$\Rightarrow R(f_0) \geq Pr[N(S) \geq m/2 \mid X = X_A] + Pr[N(S) < m/2 \mid X = X_B]$

$\boxed{\therefore R(f_0) \geq 1/2\, Pr[N(S) \geq m/2 \mid X = X_A]}$

b) Show $R(f_0) > \frac{1}{4}\left[1 - \left[1 - \exp\left[\frac{-m\varepsilon^2}{1-\varepsilon^2}\right]\right]^{1/2}\right]$

- from D.16 $Pr[B \geq K] \geq Pr\left[N \geq \frac{K - mp}{mp(1-p)^{1/2}}\right]$

$\Rightarrow N(S) \geq m/2 \mid X = X_A = B(m, p_A) \geq K, \; K = m/2,$ m = even, K = int

- $Pr[N(S) \geq m/2 \mid X = X_A] \geq Pr\left[N \geq \left(m/2 - m(1/2 - \varepsilon/2)\right)\left(m(1/2 - \varepsilon/2)(1/2 + \varepsilon/2)\right)^{-1/2}\right]$

$\geq Pr\left[N \geq \left(\frac{m\varepsilon}{2}\right)\left(\frac{m}{4}(1 - \varepsilon^2)\right)^{-1/2}\right]$

→ next page

3.19 page 2

b) - from D.17 $\Pr[N \geq u] \geq \frac{1}{2}\left(1-\left(1-\exp[-u^2]\right)^{1/2}\right)$

$$\Rightarrow u = -\frac{m\varepsilon}{2}\left(\frac{m}{4}(1-\varepsilon^2)\right)^{-1/2}$$

- $\Pr[N(s) \geq m/2 \mid X = X_A] \geq \frac{1}{2}\left(1-\left(1-\exp\left[-\left(\frac{m\varepsilon}{2}\left(\frac{m}{4}(1-\varepsilon^2)\right)^{-1/2}\right)^2\right]\right)^{1/2}\right)$

$$\geq \frac{1}{2}\left(1-\left(1-\exp\left[-\left(\frac{m^2\varepsilon^2}{4}\left(\frac{m}{4}(1-\varepsilon^2)\right)^{-1}\right)\right]\right)^{1/2}\right)$$

$$\geq \frac{1}{2}\left(1-\left(1-\exp\left[-m\varepsilon^2/(1-\varepsilon^2)\right]\right)^{1/2}\right)$$

- $R(f_0) \geq \frac{1}{2}\Pr[N(s) \geq m/2 \mid X = X_A] \geq \frac{1}{2}\left(\frac{1}{2}\left(1-\exp\left[\frac{-m\varepsilon^2}{1-\varepsilon^2}\right]\right)^{1/2}\right)$

$$\boxed{\therefore R(f_0) > \frac{1}{4}\left(1-\left(1-\exp\left[-\frac{m\varepsilon^2}{1-\varepsilon^2}\right]\right)^{1/2}\right)}$$

c) Show $R(f_0) > \frac{1}{4}\left[1-\left[1-\exp\left[\frac{-2(m/2)\varepsilon^2}{1-\varepsilon^2}\right]\right]^{1/2}\right]$

for m = odd or even

- m = odd $\Rightarrow$ $m+1$ substitution in (a)

$R(f_0) \geq \frac{1}{2}\Pr[N(s) \geq \frac{m+1}{2} \mid X = X_A]$ still holds

$$\therefore R(f_0) \geq \frac{1}{4}\left(1-\left(1-\exp\left[-\frac{2(m/2)\varepsilon^2}{1-\varepsilon^2}\right]\right)^{1/2}\right)$$ for m = even or odd

d) $\delta \geq R(f_0)$

$$\delta \geq \frac{1}{4}\left(1-\left(1-\exp\left[\frac{-2\,m/2\,\varepsilon^2}{1-\varepsilon^2}\right]\right)^{1/2}\right)$$

$$(1-4\delta)^2 \leq 1-\exp\left[\frac{-m\varepsilon^2}{1-\varepsilon^2}\right]$$

$$\log\left(1-(1-4\delta)^2\right) \geq -\frac{m\varepsilon^2}{1-\varepsilon^2}$$

$$\frac{1-\varepsilon^2}{\varepsilon^2}\log\left(\frac{1}{1-(1-4\delta)^2}\right) \leq m$$

as $\varepsilon \to 1$
$m \to 0$
as $\delta \to 1/4$
$m \to 0$