Reem Emad Ramadan sec.:1, B.N.:33

Assignment 2

** problem 2.16

$X = \mathbb{R} \quad , \quad X = 1\text{-}D \quad , \quad H = \left\{ h_c(X) = \text{sign}\left(\sum_{i=0}^{D} c_i \cdot x^i \right) \right\}$

$\text{prove } d_{VC}(H) = D+1 \quad \begin{cases} D+1 \text{ points shattered by } H & \text{(a)} \\ \text{no } D+2 \text{ points shattered by } H & \text{(b)} \end{cases}$

Soln:

it acts similar to positive ray behaviour

a) Assume we have D+1 points that gives solution for c_is where all values (D roots where the polynomial $\sum_{i=0}^{D} c_i \cdot x^i$ has D roots) fall in the interval betweem minimum and maximum point of D+1 points.

sign will change from left to right so for any Dichotomy of D+1 points, choose polynomial matches signs of each point

max. Dichotomy $= 2^N = 2^D$, $m_H(N) \leq 2^N$

$\therefore m_H(N) = D+1$

$\therefore$ there's D+1 points can be shattered by H

DATE

b) since we have D roots then for $D+2$ points, we will have at least two points on left most or right most that have same sign

$\therefore$ $D+2$ points can't be shattered by H

From ⓐ & ⓑ $d_{VC_H} = D+1$

** Problem 2.24

a) $\bar{g}(X) = \mathcal{E}_D[g^D(X)]$

$g^D[X]$: hypothesis g that gets min. E_{in} applying on dataset D

min. E_{in} at $\nabla E_{in} = 0$

$$E_{in}(g) = \sum_{i=1}^{2} [f(X_i) - h(X_i)]^2 = \sum_{i=1}^{2} [X_i^2 - (aX_i + b)]^2$$

$$\nabla_a E_{in}(g) = -2 \sum_{i=1}^{2} X_i (X_i^2 - aX_i - b) = 0$$

$$\nabla_b E_{in}(g) = -2 \sum_{i=1}^{2} (X_i^2 - aX_i - b) = 0$$

$\times X_1$

$$\nabla_b E_{in}(g) = X_1 [X_1^2 - aX_1 - b + X_2^2 - aX_2 - b] \quad — ②$$

$\times X_2$

$$\nabla_b E_{in}(g) = X_2 [X_1^2 - aX_1 - b + X_2^2 - aX_2 - b] \quad — ③$$

$$\nabla_a E_{in}(g) = X_1^2 - aX_1 - b + X_2^2 - aX_2 - b \quad — ①$$

PAGE
DATE

$(2) - (1) : X_1^2 - aX_1 - b = 0 \quad - (4)$

$(3) - (1) : X_2^2 - aX_2 - b = 0 \rightarrow b = X_2^2 - aX_2$

in (4) $X_1^2 - aX_1 - X_2^2 + aX_2 = 0$

$a(\cancel{X_2 - X_1}) = X_2^2 - X_1^2 = (\cancel{X_2 - X_1})(X_2 + X_1)$

$\therefore a = X_2 + X_1 \quad , \quad b = -X_1 X_2$

$\therefore g^D(X) = aX + b \rightarrow (X_2 + X_1)X - X_1 X_2$

$\bar{g}(X) = E_D[g^D(X)] = E_D[X_2 X + X_1 X - X_1 X_2]$

$= E_D[X_2 X] + E_D[X_1 X] - E_D[X_1 X_2]$

CamScanner

b)

1. get $\bar{g}(X)$

→ Fix X : [-1, 1]

→ For nSamples : # e.g. 1000

- sample 2 data pts from uniform distribution on [-1, 1]
- Compute $g^D(x)$ using a, b derived
- append $g^D(x)$ to list gDList

→ take mean of gDList

2. To compute bias, varience, out-of-sample error

→ For nDataSamples: # e.g.: 1000

- sample X from uniform distribution [-1, 1]
- calculate $\bar{g}(X)$ to generate array of values of function $g^D(X)$ evaluated at given X
- compute the varience: $E_D[(g^D(X) - \bar{g}(X))^2]$
- Use $\bar{g}(X)$ and calculate $(\bar{g}(X) - f(X))^2$ at each
- Use array of values to compute array of $(\bar{g}(X) - f(X))^2$
- get mean of resulting array

→ Take average of above calculated, we get $\underbrace{E_x[E_D(g^D(x) - \bar{g}(X))^2]}_{\text{varience}}$, $\underbrace{E_x[(\bar{g}(X) - f(X))^2]}_{\text{bias}}$,

$\underbrace{E_x[E_D(g^D(X) - f(X))^2]]}_{\text{out-of-sample error}}$

d)

$$\text{Varience} = E_X[(E_D[g^D(x)] - \bar{g}(x))^2]$$

$$= E_X[E_D[((X_1+X_2)X - X_1X_2 - \bar{g}(x))^2]]$$

$$\text{bias} = E_X[(\bar{g}(x) - f(x))^2]$$

$$E_{out} = E_X[E_D[(g^D(x) - f(x))^2]]$$

$\because X_1, X_2$ independence

$\therefore \bar{g}(X) = E_D[X_2]X + E_D[X_1]X - E_D[X_2]E_D[X_1]$

$\because$ data X follows uniform distribution $[-1, 1]$

$\therefore E_D[X_1] = E_D[X_2] = 0 \quad \therefore \bar{g}(X) = 0$

Varience $= E_X[E_D((X_1+X_2)^2X^2 + X_1^2X_2^2 - 2X_1X_2(X_1+X_2)X)]$

$= E_X[X^2E_D(X_1^2 + X_2^2 + 2X_1X_2) + E_D(X_1^2X_2^2) - 2XE_D(X_1^2X_2 + X_1X_2^2)]$

$= E_X[X^2(\frac{1}{3} + \frac{1}{3}) + (\frac{1}{3} \cdot \frac{1}{3})] = \frac{1}{3} = 0.33$

bias $= E_X[X^4] = \frac{1}{5} = 0.2$

$E_D[E_{out}(g^D)] = \text{var} + \text{bias} = \frac{1}{3} + \frac{1}{5} = \frac{8}{15} = 0.53$

e) Varience $= E_D [(g^D(x) - \overline{g}(x))^2]$

$C_i = x^2$

$= (c_{list} - \bar{g}(x))^2$

C_i's

bias $= (\bar{g}(x) - f(x))^2$

$E_{out} = E_D \, E_X [(g^D(x) - f(x))^2]$

$= E_D \, E_X [(c_{list} - x^2)^2]$

$= E_D \, E_X [c_{list}^2 \; x^2 + x^4]$

$= E_D [E_X(x^4) - 2 \, c_{list} \cdot E_X(x^2) + c_{list}^2]$

$= E_D [\frac{1}{5} - 2 \cdot c_{list} \cdot \frac{1}{3} + (c_{list})^2]$

$\because$ Uniform distribution

$\therefore E_D[X] = 0 , E_D[X^2] = \frac{1}{3} , E_D[X^4] = \frac{1}{5}$

** problem 3.12 : To prove H is projection matrix,
show that H is idempotent, symmetric & Positive semi-define

1 2

$\because H = X(X^TX)^{-1}X^T \therefore H^2 = X(X^TX)^{-1}X^TX(X^TX)^{-1}X^T$

$H^2 = X[(X^TX)^{-1}]^2X^T = H$ — 1

$H^T = [X(X^TX)^{-1}X^T]^T = X((X^TX)^T)^{-1}X^T$

$= X(X^TX)^{-1}X^T = H$ — 2

From 1 & 2 H is a projection matrix

$\therefore \hat{y} = Hy \therefore \hat{y}$ is the projection of y onto
the space spanned by X