Name: Hoda Gamal Hamouda Ismail

Section: 2

B.N. : 33

ID : 9203673

Subject : .... / /

Problem 2.16

x is one dimensional variable
for a hypothesis set : $H = \{h_c(x) = \text{sign}(\sum_{i=0}^{D} c_i x^i)\}$

prove that vc dimension of H is exactly $(D+1)$ by
(a) There are $(D+1)$ points which are shattered by H
(b) There are no $(D+2)$ points which are shattered by H

Solution :

(a) Consider $D+1$ distinct points each of $D+1$ dimensions,
Consider any dichotomy $y : (y_0 \dots y_D)^T \in \{-1, 1\}^{D+1}$
Such that $y_k = h_c(x_k) = \text{sign}(\sum_{i=0}^{D} c_i x_k^i)$

Since the $D+1$ points are distinct
so $|X| \neq 0$
so X is invertible
(vandermonde matrix)

$$X = \begin{bmatrix} 1 & x_0^1 & \dots & x_0^D \\ 1 & x_1^1 & \dots & x_1^D \\ \vdots & & & \\ 1 & x_D^1 & \dots & x_D^D \end{bmatrix}$$

by solving the system $Xc = y \rightarrow c = X^{-1} y$
there exist coefficients to satisfy any dichotomy

So H can shatter $D+1$ points
$m_H(D+1) = 2^{D+1}$, $d_{vc}(H) \geq D+1$

CamScanner المسموحة ضوئيا بـ

Date :
No:

in other words :

$\sum_{i=0}^{D} c_i x^i$ is a polynomial that has D roots
for D+1 points

- Since we can find coefficients by solving the system as we have just proved it
- such that the D roots fall into D intervals between min & max of D+1 points
- then the polynomial will change the sign with moving from the left to right of a root
- So for any dichotomy of D+1 points we can choose polynomial to match sign of each point

Ex. D = 1
2 points
$h_c(x) = c_0 + c_1 x$ x_0 x_1

we can find 1 degree polynomial to separate these 2 points in all possible dichotomies

(b) Consider $D+2$ distinct points
each of $D+1$ dimensions

So they are linearly dependent

and since the polynomial of degree D
so it can alternate signs up to D times
not $D+1$ times

which means there will be a point that
can't be controlled leads to 2 neighboring
points with same sign

So we can't correctly classify $D+2$
distinct points with any dichotomy

<u>EX.</u> $D = 1$
3 points $\quad x_0 \quad x_1 \quad x_2$
$h_c(x) = c_0 + c_1 x$

we can't find 1 degree polynomial that is
a line to separate these 3 points
to match <u>all</u> possible dichotomies

So H can't shatter $D+2$ points
$m_H(D+2) < 2^{D+2}$, $d_{vc}(H) \leq D+1$

from (a), (b) $d_{vc}(H) = D+1$

Subject : ....

/ /

Problem 2.24

$\boxed{a}$ $\overline{g(x)} = E_D[g^D(x)]$

$g^D(x)$ is a hypothesis that gets min Ein on dataset D

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

$$= \sum_{i=1}^{2} [x_i^2 - (ax_i + b)]^2$$

$$= \sum_{i=1}^{2} [x_i^2 - ax_i - b]^2$$

to have min Ein
set derivatives of Ein w.r.t a & b to 0

$$\frac{\partial E_{in}(g)}{\partial a} = -\cancel{2} \sum_{i=1}^{2} x_i [x_i^2 - ax_i - b] = 0 \longrightarrow (1)$$

$$\frac{\partial E_{in}(g)}{\partial b} = -2 \sum_{i=1}^{2} [x_i^2 - ax_i - b] = 0 \longrightarrow (2)$$

$x_1 * (2) - (1)$

$$\cancel{(x_1 - x_2)}(x_2^2 - ax_2 - b) = 0$$

$x_2 * (2) - (1)$

$$\cancel{(x_2 - x_1)}(x_1^2 - ax_1 - b) = 0$$

GHARIB

No:

we get: $x_1^2 - ax_1 - b = 0$

$x_2^2 - ax_2 - b = 0$

Solve both equations together

$x_1^2 - ax_1 = b$

$x_2^2 - ax_2 = b$

$x_1^2 - ax_1 = x_2^2 - ax_2$

$x_1^2 - x_2^2 = ax_1 - ax_2$

$\cancel{(x_1 - x_2)}(x_1 + x_2) = a\cancel{(x_1 - x_2)}$

$a = x_1 + x_2$

$b = x_1^2 - ax_1 = x_1^2 - (x_1 + x_2)x_1$

$= \cancel{x_1^2} - \cancel{x_1^2} - x_1 x_2$

$b = -x_1 x_2$

we get: $a = x_1 + x_2$

$b = -x_1 x_2$

$\overline{g}(x) = E_D[g^D(x)]$

$= E_D[ax + b]$

$= E_D[(x_1 + x_2)x - x_1 x_2]$

$= E_D[x_1 x + x_2 x - x_1 x_2]$

$= E_D[x_1 x] + E_D[x_2 x] - E_D[x_1 x_2]$

$= E_D[x_1]x + E_D[x_2]x$

$- E_D[x_1]E_D[x_2]$ — x_1, x_2 independent

$E_D[x_1] = ?$

x_1 is uniformaly distributed in the interval $[-1, 1]$

$$E_D[x] = \frac{1}{2}\int_{-1}^{1} x\,dx = \frac{1}{2}\left[\frac{x^2}{2}\right]_{-1}^{1} = 0$$

Same for x_2

$\overline{g}(x) = 0$

Date :
No:

d

$$\text{Variance} = E_x\left[E_D\left[\left(g^D(x) - \overline{g(x)}\right)^2\right]\right]$$

$$= E_x\left[E_D\left[\left((x_1+x_2)x - x_1x_2 - 0\right)^2\right]\right]$$

$$= E_x\left[E_D\left[(x_1+x_2)^2x^2 - 2(x_1+x_2)x\,x_1x_2 + x_1^2x_2^2\right]\right]$$

$$= E_x\left[E_D\left[(x_1+x_2)^2\right]x^2 - 2E_D\left[(x_1+x_2)x_1x_2\right]x \right.$$

$$\left. + E_D\left[x_1^2x_2^2\right]\right]$$

$$= E_x\left[x^2E_D[x_1^2] + x^2E_D[x_2^2] + 2E_D[x_1]E_D[x_2]x^2\right.$$

$$- 2E_D[x_1^2]E_D[x_2] - 2E_D[x_2^2]E_D[x_1]$$

$$\left. + E_D[x_1^2]E_D[x_2^2]\right]$$

$$E_D[x_1^2] = ?$$

$$E_D = \frac{1}{2}\int_{-1}^{1} x^2\,dx = \frac{1}{2}\left[\frac{x^3}{3}\right]_{-1}^{1} = \frac{1}{3}$$

$$\text{Variance} = E_x\left[x^2\left(\frac{1}{3}+\frac{1}{3}+0\right) - 0 - 0 + \frac{1}{3}\times\frac{1}{3}\right]$$

$$= E_x\left[\frac{2}{3}x^2 + \frac{1}{3}\times\frac{1}{3}\right] = \frac{2}{3}\times\frac{1}{3} + \frac{1}{3}\times\frac{1}{3} = \boxed{\frac{1}{3}}$$

Subject : ....

$$\text{Bias} = E_x[(\overline{g(x)} - f(x))^2]$$

$$= E_x[(0 - x^2)^2] = E_x[x^4]$$

$E_x[x^4] = ?$

$$E_x[x^4] = \frac{1}{2}\int_{-1}^{1} x^4\,dx = \frac{1}{2}\left[\frac{x^5}{5}\right]_{-1}^{1} = \frac{1}{5}$$

$$\text{Bias} = \boxed{\frac{1}{5}}$$

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

$$= E_x[E_D[(g^D(x) - f(x))^2]]$$

$$= E_x[E_D[(g^D(x))^2] - 2E_D[g^D(x)]f(x) + f(x)^2]$$

$$= E_x[E_D[(g^D(x))^2] - 2\overline{g(x)}f(x) + f(x)^2]$$

$$= E_x[E_D[(g^D(x))^2] - \overline{g(x)}^2 + \overline{g(x)}^2 - 2\overline{g(x)}f(x) + f(x)^2]$$

$$= E_x[E_D[(g^D(x) - \overline{g(x)})^2] + (\overline{g(x)} - f(x))^2]$$

$$= E_x[\text{Variance} + \text{Bias}] = \frac{1}{3} + \frac{1}{5} = \frac{8}{15}$$