Cairo University
Faculty of Engineering

Department of Computer Engineering

Assignment 1
Report
Solutions for Handwritten Problems

Student Info

Name	Abdelrahman Hamdy Ahmed Mekhaimar
Section	1
BN	38
ID	9202833

Date: 19/04/2024

Submitted to: Eng. Mohammed Shawky

ML Assignment 2

2.16) $x \rightarrow$ 1-D variable

Hypothesis Set $\Rightarrow$ $H = \left\{ h_c(x) = \text{sign}\left(\sum_{i=0}^{D} c_i x^i \right) \right\}$

Prove that the vc dimension of H is exactly $(D+1)$ by showing that

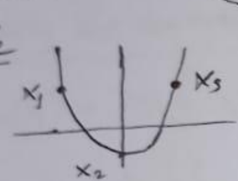
(a) There are $(D+1)$ points which are shattered by H

The summation $\sum_{i=0}^{D}$ implies that the total points needed to solve the polynomial are D+1 since it's guaranteed that it'll pass through them all. In other words, the polynomial $\sum_{i=0}^{D} c_i x^i$ has D roots. For any Dichotomy of D+1 points (assigning signs to the chosen D+1 points), we can solve the System of equations to get the coefficients that will yield the signs set before. $\therefore$ H shatters D+1 points

$m_H(N) \leq 2^N$ (Max)

$m_H(N) = D+1$

example



$\left.\begin{array}{l} x_1 \rightarrow +ve \\ x_2 \rightarrow -ve \\ x_3 \rightarrow +ve \end{array}\right\}$ Solve & get coefficients

(b) For D+2 points, we only have D roots therefore they'll only insight us on D+1 points. In this case, we will have atleast 2 points on the leftmost or rightmost regions that'll have the same sign. therefore we won't be able to find a straight line that separates them thus it won't be counted as shattering. This will prevent us from generating all dichotomies for D+2 points due to the fact that we can only control the signs of D+1 points.

Conclusion: 2 neighboring Points having the same sign will prevent us from placing all orientations for the growth function to be D+2

ex
D=1

From (a) & (b)

vc dimension of H is exactly D+1

$$\boxed{d_{vc_H} = D+1}$$

2.24 $x \to$ Uniformly Distributed $[-1, 1]$

Dataset consists of 2 points $\{x_1, x_2\}$ $D = \{(x_1, x_1^2), (x_2, x_2^2)\}$

Target function $f(x) = x^2$

The learning algorithm returns the line fitting these 2 points as g

$$H \to h(x) = ax + b$$

(a) Give the analytic expression of the avg function $\bar{g}(x)$

$\bar{g}(x) = E_D[g^D(x)]$ To get minimum $E_{in} \to \nabla E_{in} = 0$

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

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

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

Expanding the summation

$$\nabla_a E_{in}(g) = \cancel{-2}[x_1^3 - ax_1^2 - bx_1 + x_2^3 - ax_2^2 - bx_2] = 0$$

$$\nabla_b E_{in}(g) = \cancel{-2}[x_1^2 - ax_1 - b + x_2^2 - ax_2 - b] = 0$$

$$x_1 \cdot \nabla_b E_{in}(g) = x_1^3 - ax_1^2 - bx_1 + x_1 x_2^2 - a x_2 x_1 - bx_1 = 0 \quad \cdots (1$$

$$x_2 \cdot \nabla_b E_{in}(g) = x_2 x_1^2 - ax_1 x_2 - bx_2 + x_2^3 - ax_2^2 - bx_2 = 0 \quad \cdots (2$$

$$\nabla_a E_{in}(g) = x_1^2 - ax_1 - b + x_2^2 - ax_2 - b \quad \cdots (3$$

$(1 - (2 \to x_1^2 - ax_1 - b = 0 \quad \cdots (4$

$(2 - (1 \to x_2^2 - ax_2 - b = 0 \longrightarrow b = x_2^2 - ax_2$

(4) $x_1^2 - ax_1 - x_2^2 + ax_2 = 0$

$$a(x_2 - x_1) = x_2^2 - x_1^2 = (x_2 - x_1)(x_2 + x_1)$$

$$a = x_2 + x_1 \quad \& \quad b = -x_1 x_2$$

$$g^D(x) = ax + b \to (x_2 + x_1)x - x_1 x_2$$

$$\bar{g}(x) = E_D[g^D(x)] = E_D[x_2 x + x_1 x - x_1 x_2]$$

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

$$\bar{g}(x) = \underbrace{E_D[x_1]x + E_D[x_2]x - E_D[x_1]E_D[x_2]}_{\text{valid due to the independence between } x_1 \,\&\, x_2}$$

$\hookrightarrow$ uniform Distribution $[-1, 1]$

$\Rightarrow E(x_1) = 0$

$\& \ E_D(x_2) = 0$

$E_D(x) = 0$

$$\bar{g}(x) = 0 \cdot x + 0 \cdot x - 0 \cdot 0 = \boxed{0}$$

2.24) a) Another way is to formulate the expectation across Dataset D as the entire interval -1 to 1 using the double integral for x_1 & x_2

$g(x) \rightarrow ax+b$

$a \rightarrow$ slope

$b \rightarrow$ y-intercept

$$\text{Slope} = \frac{y_1 - y_2}{x_1 - x_2}$$

y-intercept (b)

$$= \frac{x_1 y_2 - x_2 y_1}{x_1 - x_2}$$

$$\bar{g}(x) = E_D[g(x)]$$

$$= E_D\left[\frac{y_1 - y_2}{x_1 - x_2} x + \frac{x_1 y_2 - x_2 y_1}{x_1 - x_2}\right]$$

$$= \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} \frac{x_1^2 - x_2^2}{x_1 - x_2}\, dx_1 dx_2 \cdot x$$

$$+ \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} \frac{x_1 x_2^2 - x_2 x_1^2}{x_1 - x_2}\, dx_1 dx_2$$

$$= \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} (x_1 + x_2)\, dx_1 dx_2 \cdot x - \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} (x_1 x_2)\, dx_1 dx_2$$

$$= \frac{1}{4} \cdot 0 - \frac{1}{4} \cdot 0 = \boxed{0}$$ ✓

2.24) (d) Compute analytically what E_{out}, bias & var should be.

$$E_{out} = E_x[(g(x) - f(x))^2] = E_x[(ax + b - x^2)^2]$$

$$= E_x[x^4] - 2aE_x[x^3] + (a^2 - 2b)E_x[x^2] + 2abE_x[x] + b^2$$

$$= \frac{1}{2}\int_{-1}^{1} x^4 dx - 2a\frac{1}{2}\int_{-1}^{1} x^3 dx + (a^2-2b)\cdot\frac{1}{2}\int_{-1}^{1} x^2 dx + ab\int_{-1}^{1} x\,dx + b^2$$

$$= \frac{1}{5} + \frac{(a^2-2b)}{3} + b^2$$

Expectation with respect to $\mathcal{D}$ (Replace a by x_1+x_2 & b by $-x_1x_2$)

$$E_{\mathcal{D}}[E_{out}] = \frac{1}{5} + \frac{1}{3}E_D[(x_1+x_2)^2 + 2x_1x_2] + E_D[x_1^2x_2^2]$$

$$= \frac{1}{5} + \frac{1}{3}\cdot\frac{1}{4}\int_{-1}^{1}\int_{-1}^{1}(x_1^2 + x_2^2 + 4x_1x_2)\,dx_1dx_2 + \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} x_1^2x_2^2\,dx_1dx_2$$

$$= \frac{1}{5} + \frac{1}{3}\cdot\frac{1}{4}\cdot\frac{8}{3} + \frac{1}{4}\cdot\frac{4}{9} = \boxed{\frac{8}{15}}$$

· Bias $(x) = (\bar{g}(x) - f(x))^2 = f(x)^2 = x^4$

$$\rightarrow \text{bias} = E_x[x^4] = \frac{1}{2}\int_{-1}^{1} x^4 dx = \frac{1}{5}$$

· $\text{var}(x) = E_D[(g(x) - \bar{g}(x))^2] = E_D[a^2x^2 + 2abx + b^2]$

$$= E_D[a^2]\cdot x^2 + 2E_D[ab]\cdot x + E_D[b^2]$$

$$= E_D[(x_1+x_2)^2]\cdot x^2 - 2E_D[(x_1+x_2)x_1x_2]\cdot x + E_D[x_1^2x_2^2]$$

$$= E_D[x_1^2 + 2x_1x_2 + x_2^2]\cdot x^2 - 2E_D[x_1^2x_2 + x_1x_2^2]\cdot x + E_D[x_1^2x_2^2]$$

$$= \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1}(x_1^2 + 2x_1x_2 + x_2^2)\,dx_1dx_2\cdot x^2 - \frac{2}{4}\int_{-1}^{1}\int_{-1}^{1}(x_1^2x_2 + x_1x_2^2)\,dx_1dx_2\cdot x$$

$$+ \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} x_1^2x_2^2\,dx_1dx_2 = \frac{1}{4}\left(\frac{4}{3} + 0 + \frac{4}{3}\right)\cdot x^2 - 0\cdot x + \frac{1}{4}\cdot\frac{4}{9}$$

$$= \frac{2}{3}\cdot x^2 + \frac{1}{9}$$

$$\text{var} = E_x\left[\frac{2}{3}x^2 + \frac{1}{9}\right] = \frac{2}{3}\cdot\frac{1}{2}\int_{-1}^{1} x^2 dx + \frac{1}{9} = \boxed{\frac{1}{3}}$$