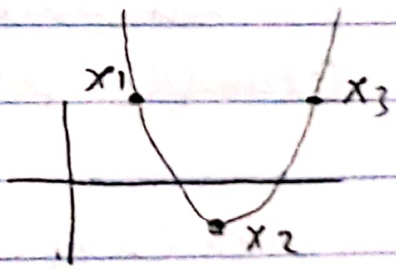
2.16

(a) To solve the polynomial $\sum_{i=0}^{D} c_i x^i$ we need D+1 points. The resulting polynomial will pass through all of these points. So we will choose D+1 points and assign the y value with the sign we need (+1 / -1). After solving the system we will get the coeffecients that will produce the intended signs on the D+1 points. so H shatter D+1 points.



• Example: for D = 2

Points: $(x_1, +1)$

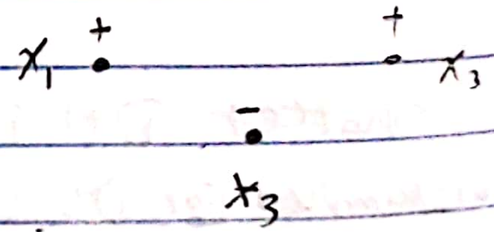
$(x_2, -1)$ ⟹ solve and get c_0, c_1, c_3.

$(x_3, +1$

we also choose any positive number other than (1) or any negtive number other than (-1).

(b) Consider the same approch as in the previous step, we can only control the sign of D+1 points, we trying to check D+2, we will notic that there is extra point which its sign can not be controlled. The case will cause that two neighboring points will have the same sign which will prevent us from generating all dichotomies

Example, let D=1



This case cannot be solved by a straigh line (no other orientation we let it be solved)
so H don't shatter D+2 point

*From (a), (b) the VC dimension of H is exactly (D+1).

$\boxed{2.24}$

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

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

set the derivatives with respect to a and b to zero

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

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

$x_1 * ① - x_2 * ②$

$x_1^2 - ax_1 - b = 0$

$x_2^2 - ax_2 - b = 0$

solve

$a = x_1 + x_2$

$b = -x_1 x_2$

$$\therefore \overset{D}{g}(x) = (\overbrace{x_1 + x_2}^{a})\, x - \overbrace{x_1 x_2}^{b}$$

$$\bar{g}(x) = E_D\left[\overset{D}{g}(x)\right]$$

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

$$= E_D\left[x_1 x\right] + E_D\left[x_2 x\right] - E_D\left[x_1 x_2\right]$$

$$= \underline{E_D\left[x_1\right] x + E_D\left[x_2\right] x - E_D\left[x_1\right] E_D\left[x_2\right]}$$

due to independene
between x_1 and x_2

As the data follow a uniform distribution
betwen $-1, 1$: $E(x_1) = 0$ & $E_D(x_2) = 0$ so
$E_D^D(x) = 0$

$$\therefore \bar{g}(x) = 0 * x + 0 * x - 0 * 0$$

$$= 0$$

(d)

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

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

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

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

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

$\left. - 2xE_D\left(x_1^2x_2 + x_1x_2^2\right)\right]$

As the data follow an uniform distribution between -1, -1:

$E(x) = 0$

$E(x^2) = \frac{1}{3}$

$E(x^3) = 0$

$E(x^4) = \frac{1}{5}$

$= Var = E_x\left[X^2\left(\frac{1}{3}+\frac{1}{3}\right) + \left(\frac{1}{3}\cdot\frac{1}{3}\right)\right]$

other term with X_1 & X_2 yields zero

$$= E_x\left(\frac{2}{3}X^2 + \frac{1}{9}\right)$$

$$= \frac{2}{3} \times \frac{1}{3} + \frac{1}{9} = \frac{1}{3}$$

$$Bias = E_x\left[\left(\bar{g}(x) - f(x)\right)^2\right]$$

$$= E_x\left[\left(0 - x^2\right)^2\right]$$

$$= E_x\left[X^4\right] \longrightarrow \text{from prev page}$$

$$= \frac{1}{5}$$

$$E_{out} = Bias + Variance$$

$$= \frac{1}{3} + \frac{1}{5} = \frac{8}{15}$$

$$E_x[X] = \frac{1}{2}\int_{-1}^{1} X \, dX = \frac{1}{2} \times \frac{1}{2}\left[1^2 - (-1)^2\right] = 0$$

$$E_x[X) = \frac{1}{2}\int_{-1}^{1} X^2 \, d \cdot X = \frac{1}{2} \times \frac{1}{3}\left[1^3 - (-1)^3\right] = \frac{1}{3}$$

$$E_x[X^3) = \frac{1}{2}\int_{-1}^{1} X^3 \, dX = \frac{1}{2} \times \frac{1}{4}\left[1^4 - (-1)^4\right] = 0$$

$$E_x[X^4] = \frac{1}{2}\int_{-1}^{1} X^4 \, dX = \frac{1}{2} \times \frac{1}{5}\left[1^5 - (-1)^5\right] = \frac{1}{5}$$

$\boxed{3.12}$

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

let $X^TX = A$

$$H^2 = X\underbrace{A^{-1}A}_{I}A^{-1}X^T$$

$$= XA^{-1}X^T$$

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

$\therefore \hat{y}$ is the projection of y onto the column space of X.