Savinay Shukla - Midterm

Q1) We are given the following emperical Risk

$$R(\theta) = \frac{1}{2N} \| y - X\theta \|^2 + \theta^T\theta + a^T\theta$$

↳ Simplifying

$$R(\theta) = \frac{1}{2N} \left((y - X\theta)^T (y - X\theta) \right) + \theta^T\theta + a^T\theta$$

$$= \frac{1}{2N} \left(y^T y - 2y^T X\theta + \theta^T X^T X\theta \right) + \theta^T\theta + a^T\theta$$

Taking the partial derivative w.r.t θ :-

$$\frac{\partial R(\theta)}{\partial \theta} = \frac{1}{2N} \left(-2y^T X + 2\theta^T X^T X \right) + 2\theta^T + a^T$$

$$= -\frac{y^T X}{N} + \frac{\theta^T (X^T X)}{N} + 2\theta^T + a^T$$

$$= -\frac{y^T X}{N} + \theta^T \left(\frac{X^T X}{N} + 2I \right) + a^T$$

Putting $\frac{\partial R(\theta)}{\partial \theta} = 0$

$$↳ -\frac{X^T y}{N} + \left(\frac{X^T X}{N} + 2I \right)^T \theta + a = 0$$

$$↳ \left(\frac{X^T X}{N} + 2I \right)^T \theta = -a + \frac{X^T y}{N}$$

$$\theta^* = \left[\frac{X^T X}{N} + 2I \right]^{-1} \left(-a + \frac{X^T Y}{N} \right) //$$

Q2) Given $w = xy^2$; $x = r\cos\theta$; $y = r\sin\theta$

Sol^n:-

$$\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x} \cdot \frac{\partial x}{\partial r} + \frac{\partial w}{\partial y} \cdot \frac{\partial y}{\partial r}$$

$$= \frac{\partial (xy^2)}{\partial x} \cdot \frac{\partial (r\cos\theta)}{\partial r} + \frac{\partial (xy^2)}{\partial y} \cdot \frac{\partial (r\sin\theta)}{\partial r}$$

$$= y^2\cos\theta + 2xy\sin\theta$$

Given $r = 2$ and $\theta = \pi/4$:

$$\frac{\partial w}{\partial r} = r^2\sin^2\theta\cos\theta + 2r^2\sin^2\theta\cos\theta$$

$$= \left(4 \times \frac{1}{2} \times \frac{1}{\sqrt{2}}\right) + \left(2 \times 4 \times \frac{1}{2} \times \frac{1}{\sqrt{2}}\right)$$

$\because \sin\frac{\pi}{4} = \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}}$

$$= \frac{2}{\sqrt{2}} + \frac{4}{\sqrt{2}}$$

$$\frac{\partial w}{\partial r} = 3\sqrt{2} \quad //$$

Q3) Given $f(x|\theta) = \frac{1}{\theta} x^{(1-\theta)/\theta}$

for $0 < x < 1$, $0 < \theta < \infty$

$$L(\theta) = \prod_i^n f(x_i|\theta)$$

$$= \prod_i^n \left(\frac{1}{\theta}\right) x_i^{(1-\theta)/\theta}$$

$$= \prod_i^n \left(\frac{1}{\theta}\right) x_i^{1/\theta - 1}$$

$$= \left(\frac{1}{\theta}\right)^n \prod_i^n \left[x_i^{(1/\theta - 1)} \right]$$

Taking Log:

$$\text{Log}(L) = \log \left[\left(\frac{1}{\theta}\right)^n \prod_i \left[x_i^{(1/\theta - 1)} \right] \right]$$

$$= n \log\left(\frac{1}{\theta}\right) + \left(\frac{1}{\theta} - 1\right) \sum_i^n \log x_i$$

diff w.r.t θ and minimize :-

$$\frac{\partial \log L}{\partial \theta} = -n\left(\frac{1}{\theta}\right) - \frac{1}{\theta^2} \sum_i^n \log x_i = 0$$

$$\therefore \quad -\frac{n}{\theta} = \frac{1}{\theta^2} \sum_i^n \log x_i$$

$\therefore \; -n\theta = \sum_{i}^{n} \log x_i$

$\hookrightarrow \theta = -\frac{1}{n} \sum_{i}^{n} \log(x_i)$

Given $x_1 = 0.1$, $x_2 = 0.22$, $x_3 = 0.54$ & $x_4 = 0.36$

Here $n = 4$

& $\sum_{i}^{n} \log(x_i) = \log(x_1) + \log(x_2) + \log(x_3) + \log(x_4)$

$$\therefore \; \theta = -\frac{\sum_{i}^{n} \log(x_i)}{N}$$

$$= 1.36$$

Q4) $f(x,y) = x \cdot e^y$

Constraint $\rightarrow x^2+y^2 = 2$

$$\mathcal{L}(x,y) = x \cdot e^y - \lambda(x^2+y^2-2) = 0$$

$$\left.\begin{aligned} \frac{\partial L}{\partial x} &= e^y - 2\lambda x \\ \frac{\partial L}{y} &= xe^y - 2\lambda y \end{aligned}\right\} \text{Set these values to 0}$$

$\therefore e^y = 2\lambda x$ & $x \cdot e^y = 2\lambda y$

$$\therefore \frac{e^y}{2x} = \frac{xe^y}{2y} \rightarrow \frac{1}{x} = \frac{x}{y}$$

$$\rightarrow y = x^2$$

Substituting back in the constraint :-

$$x^2 + x^2 = 2 \rightarrow 2x^2 = 2$$

$$\rightarrow x = \pm 1$$

$$\therefore y = x^2 \rightarrow x^2+y^2-2=0$$

$$\hookrightarrow y^2+y-2=0$$

$$\hookrightarrow y = -2 \text{ or } -1$$

It cant -2 because $x^2 = y$!

Square value can't be negative

$\therefore$ $y = 1$ and $x = \pm 1$.

$$\lambda = \frac{e^y}{2x} = \frac{e^1}{2} = \frac{e}{2}.$$

for Max :- $x = 1$, $y = 1$ & $\lambda = e/2$.

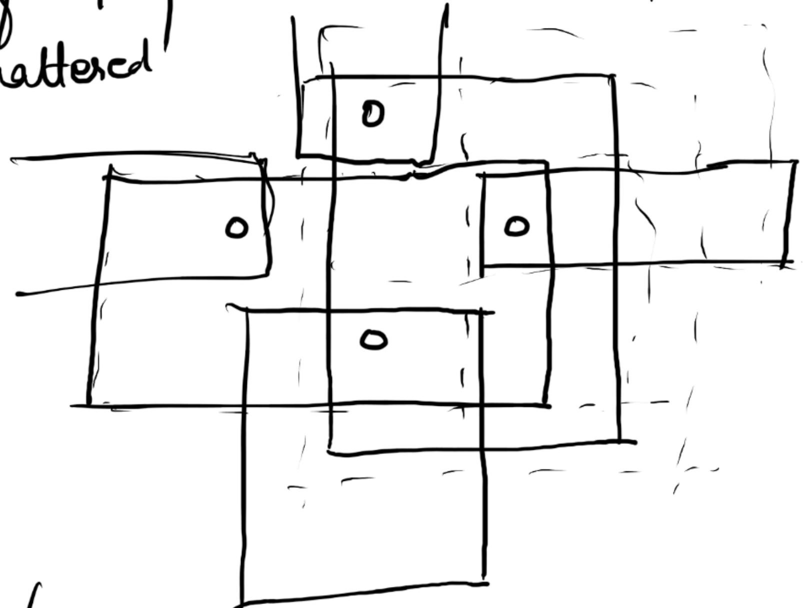
$$f(x) = xe^y - \lambda(x^2 + y^2 - 2)$$
$$= e - \frac{e}{2}(0)$$
$$f(x) = e \text{ for Max}$$

for Min :- $x = -1$, $y = 1$ & $\lambda = -\frac{e}{2}$

$$f(x) = xe^y - \lambda(x^2 + y^2 - 2)$$
$$= -e - \left(-\frac{e}{2}\right)(0)$$
$$f(x) = -e$$

$\hookrightarrow$ for Min.

Q5) For a family of 2D align rectangles we can think of an arrangment of 4 points where all points are shattered



4 points can be shattered.

But for 5, there will always be a point surrounded by extreme points.

∴ Not possible to shatter 5 points

∴ VC Dimension = 4

Q6)

Boxes	Apple	Orange
1	n_1	r_1
2	n_2	r_2
3	⋮	⋮
⋮		
m	n_1	r_1

i) Prob of picking box j

$P(B_j) = \frac{1}{m}$

ii) Prob of Selecting ~~any~~ one apple from any box

↳ $P(A | B_t) = n_t / n_t + r_t$

(iii) Prob of select one apple from exactly box j

↳ $$P(A/B_j) = \frac{P(B_j)\, P(A/B_j)}{\sum_{K}^{m} P(B_K)\, P(A/B_K)}$$

$$= \frac{\frac{1}{m} \times \frac{n_j}{n_j + r_j}}{\sum_{K}^{m} P(B_K)\, P(A/B_K)}$$

$$P(A/B_j) = \frac{n_j}{m(n_j + r_j) \times \sum_{K=1}^{m} P(B_K)\, P(A/B_K)}$$