UNIVERSITY OF MINNESOTA
Driven to Discover℠

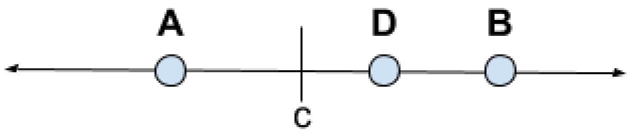
Homework #1

CSCI 5521
Professor Kuang

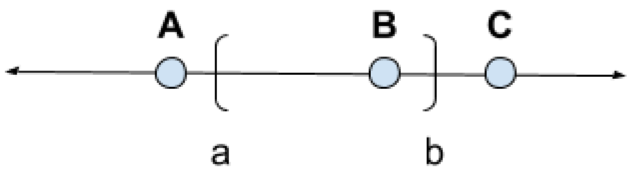
Christopher White
February 28, 2019

Question 1

A. The set ℝ is a one dimensional space. It can be represented by a line. If a threshold of c is used, the line is split into two pieces. If one point, A, is less than c and one point, B, is greater than c, then this function shatters 2 points. If a third point, D, is placed anywhere on the line, the threshold cannot shatter the 3 points because not all 2^3 possible labelings can be separated, e.g. AB cannot be separated from D. Therefore, the VC dimension, d_c, is **2**.



B. The set ℝ is a one dimensional space. It can be represented by a line. If an interval, [a,b], is applied it can be split into three pieces. If two points, A,B, are placed on the line, the interval can select both, one, or none of the points. Therefore, the interval can shatter 2 points. However, if a third point, C, is added the interval cannot separate all 2^3 possible labelings so it cannot shatter 3 points, e.g. AC cannot be separated from B. Therefore, the VC dimension, d_I, is **2**. θ



Question 2

A. $l(\theta|x) \;=\prod_{t=1}^{n} \frac{1}{\theta} exp(-\frac{x^t}{\theta})$

$$L(\theta|x) \;= log\; l(\theta|x) \;= \sum_{t=1}^{n} log(\frac{1}{\theta} exp(-\frac{x^t}{\theta})) \;=\; \sum_{t=1}^{n} - log(\theta) \;- (\frac{x^t}{\theta}) \;=\; -n\; log(\theta) \;-\; \frac{\sum_{t=1}^{n} x^t}{\theta}$$

$$\frac{d}{d\theta} L(\theta|x) \;= 0 = \frac{d}{d\theta}\left[- n\; log(\theta) \;-\; \frac{\sum_{t=1}^{n} x^t}{\theta} \right] \;=\; -\frac{n}{\theta} + \frac{\sum x^t}{\theta^2}$$

$$-\frac{n}{\theta} + \frac{\sum x^t}{\theta^2} \;= 0 \Rightarrow \frac{n}{\theta} = \frac{\sum x^t}{\theta^2} \Rightarrow \theta = \frac{\sum_{t=1}^{n} x^t}{n}$$

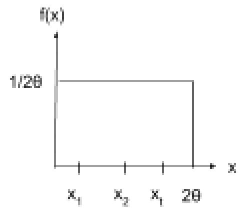
B. $l(\theta|x) \;=\prod_{t=1}^{n} 2\theta\; \{x^t\}^{2\theta-1}$

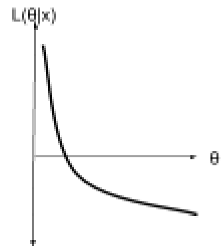
$$L(\theta|x) \;= log\; l(\theta|x) \;= \sum_{t=1}^{n} log\left(2\theta\{x^t\}^{\;2\theta-1}\right) = \sum_{t=1}^{n} log(2\theta) + (2\theta-1)\; log(x^t) = n\; log(2\theta) + (2\theta-1) \sum_{t=1}^{n} log(x^t)$$

$$\frac{d}{d\theta} L(\theta|x) \;= 0 = \frac{n}{\theta} + 2 \sum_{t=1}^{n} log(x^t) \Rightarrow \theta = \;-\;\frac{n}{2\sum_{t=1}^{n} log(x^t)}$$

C. $l(\theta|x) = \prod_{t=1}^{n} \frac{1}{2\theta} \Rightarrow L(\theta|x) \;=\; -\sum_{t=1}^{n}\; log(2\theta) = \;-n\; log(2\theta)$

$$\frac{d}{d\theta} L(\theta|x) \;= 0 =- \frac{n}{\theta}$$





The likelihood function is not minimal or maximal so the typical approach doesn't work. The following relationship does hold.

$$x_1, x_2, ..., x_t \leq 2\theta \;\Rightarrow \widehat{\theta} \geq max(x_1, x_2, ..., x_t)/2$$

Question 3

A. $P(C_i|x) = \frac{p(x|C_i)P(C_i)}{p(x)} = \frac{p(x|C_i)P(C_i)}{\sum_{k=1}^{K} p(x|C_k)P(C_k)}$

$P(x|C_i) = p_i^{1-x}(1-p_i)^x$

$P(C_1|x=0) = \frac{p_1 P(C_1)}{p_1 P(C_1)+p_2 P(C_2)}$ $\quad P(C_2|x=0) = 1 - P(C_1|x=0) = \frac{p_2 P(C_2)}{p_1 P(C_1)+p_2 P(C_2)}$

$P(C_1|x=1) = \frac{(1-p_1)P(C_1)}{(1-p_1)P(C_1)+(1-p_2)P(C_2)}$ $\quad P(C_2|x=1) = 1 - P(C_1|x=1) = \frac{(1-p_2)P(C_2)}{(1-p_1)P(C_1)+(1-p_2)P(C_2)}$

Rule:

Choose C_1 if :

$$\left(\frac{p_1 P(C_1)}{p_1 P(C_1)+p_2 P(C_2)}\right)^{1-x}\left(\frac{(1-p_1)P(C_1)}{(1-p_1)P(C_1)+(1-p_2)P(C_2)}\right)^{x} \geq \left(\frac{p_2 P(C_2)}{p_1 P(C_1)+p_2 P(C_2)}\right)^{1-x}\left(\frac{(1-p_2)P(C_2)}{(1-p_1)P(C_1)+(1-p_2)P(C_2)}\right)^{x}$$

Otherwise, Choose C_2

B. $P(C_i|x_j) = \frac{p_{ij}^{1-x}(1-p_{ij})^x P(C_i)}{\sum_{k=1}^{K} p_{kj}^{1-x}(1-p_{kj})^x P(C_k)}$ $\quad P(x\,|C_i) = \prod_{j=1}^{D} p_{ij}^{1-x_j}\,(1-p_{ij})^{x_j}$

$P(C_i|x\,) = \frac{P(x|C_i)\,P(C_i)}{\sum_{k=1}^{K} P(x|C_k)\,P(C_k)}$

$P(C_i|x\,) = \frac{P(C_i)\,\prod_{j=1}^{D} p_{ij}^{1-x_j}\,(1-p_{ij})^{x_j}}{\sum_{k=1}^{K} P(C_k)\,\prod_{j=1}^{D} p_{kj}^{1-x_j}\,(1-p_{kj})^{x_j}}$

Rule:

Choose C_1 if:

$$P(C_1)\,\prod_{j=1}^{D} p_{1j}^{1-x_j}\,(1-p_{1j})^{x_j} \geq P(C_2)\,\prod_{j=1}^{D} p_{2j}^{1-x_j}\,(1-p_{2j})^{x_j}$$

Otherwise, Choose C_2

C. $P(C_1|x\,) = \frac{P(C_1)\,\prod_{j=1}^{2} p_{1j}^{1-x_j}\,(1-p_{1j})^{x_j}}{\sum_{k=1}^{2} P(C_k)\,\prod_{j=1}^{2} p_{kj}^{1-x_j}\,(1-p_{kj})^{x_j}}$ $\quad P(C_2|x\,) = \frac{P(C_2)\,\prod_{j=1}^{2} p_{1j}^{1-x_j}\,(1-p_{1j})^{x_j}}{\sum_{k=1}^{2} P(C_k)\,\prod_{j=1}^{2} p_{kj}^{1-x_j}\,(1-p_{kj})^{x_j}}$

Posterior Probabilities:

	$P(C_1 \mid x_1,x_2)$			$P(C_2 \mid x_1,x_2)$		
	$P(C_1)=0.2$	$P(C_1)=0.6$	$P(C_1)=0.8$	$P(C_2)=0.8$	$P(C_2)=0.4$	$P(C_2)=0.2$
$x_1=0,\ x_2=0$	0.0270	0.1429	0.3077	0.9730	0.8571	0.6923
$x_1=0,\ x_2=1$	0.6923	0.9310	0.9730	0.3077	0.0690	0.0270
$x_1=1,\ x_2=0$	0.0270	0.1429	0.3077	0.9730	0.8571	0.6923
$x_1=1,\ x_2=1$	0.6923	0.9310	0.9730	0.3077	0.0690	0.0270

Question 4

Validation Results

```
TABLE OF ERROR RATES

P(C1|sigma)   Error Class 1      Error Class 2

------------------------------------------

1e-05         0.54    0

9.9995e-05    0.54    0

0.0009995     0.54    0

0.0099502     0.54    0

0.095163      0.465   0.045

0.63212       0.325   0.19

0.86466       0       0.455

0.95021       0       0.46

0.98168       0       0.46

0.99326       0       0.46

0.99752       0       0.46
```

Test Results

```
TABLE OF TEST RESULTS

P(C1|sigma)  Error Class 1     Error Class 2

------------------------------------------

0.86466       0.005       0.44
```