Dolly Sidar
2019804

Assignment-2

Date: / /
Page No.:

2) a) Prior $P(\theta) = \prod_{i=1}^{d} \theta_i e^{-\sum_{i=1}^{d} \theta_i}$

For d-dimensional multivariate Bernoulli space

$P(D|\theta) = \prod_{i=1}^{n} \prod_{i=1}^{d} \theta_i^{x_i} (1-\theta_i)^{1-x_i}$

$\arg\max_\theta P(\theta) \cdot P(D|\theta) = \arg\max_\theta \prod_{i=1}^{d} \theta_i e^{-\sum_{i=1}^{d} \theta_i} \prod_{i=1}^{n} \prod_{j=1}^{d} \theta_j^{x_j} (1-\theta_j)^{1-x_j}$

$\Rightarrow \sum_{i=1}^{d} \ln\theta - \sum_{i=1}^{d} \theta_i + \sum_{i=1}^{n} \sum_{j=1}^{d} x_j \ln\theta_j + (1-x_j)\ln(1-\theta_j)$

Now, differentiating w.r.t θ_j, we get

$\Rightarrow \theta_j^{-1} - 1 + \sum_{j=1}^{n} x_j/\theta_j - \frac{(1-x_j)}{(1-\theta_j)} = 0$

$\Rightarrow \frac{1-\theta_j}{\theta_j} + \sum_{j=1}^{n} \frac{x_j(1-\theta_j) - \theta_j(1-x_j)}{\theta_j(1-\theta_j)} = 0$

$\Rightarrow \frac{1-\theta_j}{\theta_j} + \sum_{j=1}^{n} \frac{x_j - \cancel{x_j\theta_j} - \theta_j + \cancel{\theta_j x_j}}{\theta_j(1-\theta_j)} = 0$

$\Rightarrow \frac{1-\theta_j}{\theta_j} + \sum_{j=1}^{n} \frac{x_j-\theta_j}{\theta_j(1-\theta_j)} = 0$

$\Rightarrow \frac{\cancel{\theta_j}(1-\theta_j)}{\theta_j} - \theta_j(1-\theta_j) + \sum_{j=1}^{n} (x_j-\theta_j) = 0$

$\Rightarrow 1-\theta_j - \theta_j + \theta_j^2 + \sum_{j=1}^{n} x_j - n\theta_j = 0$

$\Rightarrow 1 - 2\theta_j + \theta_j^2 + \sum_{j=1}^{n} x_j - n\theta_j = 0$

$\Rightarrow \theta_j^2 - \theta_j(n+2) + 1 + \sum_{j=1}^{n} x_j = 0$

Teacher's Signature ....

Therefore,

$$\theta_{MAP} = \frac{2+n \pm \sqrt{(2+n)^2 - 4\left(\sum_{j=1}^{n} x_j + 1\right)}}{2} \quad \text{--- (i)}$$

b) Given, $x = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$

no. of samples $(n) = 4$

from eq (i)

$$\theta^1_{MAP} = \frac{2+4 \pm \sqrt{(2+4)^2 - 4\left(\sum_{j=1}^{4} x_j + 1\right)}}{2}$$

$$= \frac{6 \pm \sqrt{6^2 - 4(3+1)}}{2}$$

$$= \frac{6 \pm \sqrt{36-16}}{2} = \frac{6 \pm \sqrt{20}}{2} = \frac{2(3 \pm \sqrt{5})}{2}$$

$$\boxed{\theta^1_{MAP} = 3 \pm \sqrt{5}}$$

$$\theta^2_{MAP} = \frac{6 \pm \sqrt{6^2 - 4(3-1)}}{2}$$

$$= \frac{6 \pm \sqrt{36-8}}{2}$$

$$= \frac{6 \pm \sqrt{28}}{2} = \frac{2(3 \pm \sqrt{7})}{2}$$

$$\boxed{\theta^2_{MAP} = 3 \pm \sqrt{7}}$$

Q 1) b) We know that for a multivariate Bernoulli variable with d dimensions and N samples-

$$x_i = [x_{i1} \ x_{i2} \ \ldots \ x_{id}]^T$$
$$x = [x_1 x_2 \ldots \ldots x_N]$$
$$\theta = \{\theta_1 \ \theta_2 \ \ldots \ \theta_d\}$$

$$P(x_i|\theta) = \prod_{j=1}^{d} \theta_j^{x_{ij}} (1-\theta_j)^{1-x_{ij}}$$

$$P(x|\theta) = \prod_{i=1}^{N} \prod_{j=1}^{d} \theta_j^{x_{ij}} \cdot (1-\theta_j)^{(1-x_{ij})}$$

Taking ln on both sides we have-

$$\ln P(x|\theta) = \ln \left\{ \prod_{i=1}^{N} \prod_{j=d}^{d} \theta_j^{x_{ij}} (1-\theta_j)^{(1-x_{ij})} \right\}$$
$$= \sum_{i=1}^{N} \sum_{j=1}^{d} \ln \left\{ \theta_j^{x_{ij}} (1-\theta_j)^{(1-x_{ij})} \right\}$$
$$= \sum_{i=1}^{N} \sum_{j=1}^{d} \{ x_{ij} \ln \theta_j + (1-x_j) \ln (1-\theta_j) \}$$

log likehihood function $F(\hat{\theta})$ is

$$F(\theta) = \sum_{i=1}^{N} \sum_{j=1}^{d} \{ x_{ij} \ln \theta_j + (1-x_{ij}) \ln (1-\theta_j) \}$$

$$\frac{\partial F(\theta)}{\partial \theta_j} = \sum_{i=1}^{N} \left\{ \frac{x_{ij}}{\theta_j} - \left(\frac{1-x_{ij}}{1-\theta_j} \right) \right\}$$

Equating above to 0

$$\sum_{i=1}^{N} \left\{ \frac{x_{ij}(1-\theta_j) - (1-x_j)\theta_j}{\theta_j(1-\theta_j)} \right\} = 0$$

Date: / /
Page No.:

$$\sum_{i=1}^{N} \{ x_{ij} - x_{ij}/\theta_j - \theta_j + x_{ij}/\theta_j \} = 0$$

$$\sum_{i=1}^{N} x_{ij} - N\theta_j = 0$$

$$N\theta_j = \sum_{i=1}^{N} x_{ij}$$

$$\boxed{\theta_j = \frac{1}{N} \sum_{i=1}^{N} x_{ij}} \leftarrow \text{MLE}$$

Using the above formula all θ_j $j = 1, 2 \dots d$ can be calculated to come up with θ.

Observation in Graphs for MLE v/s n where $n = 1, 2, 3 \dots 50$ is

For low values of n, MLE gives incorrect estimates. However as the size of the sample grows i.e n increases, the estimates of θ comes closer and closer to the real value of the mean

e) Let x be a sample of multivariate Bernoulli with d dimensions

$$x = [x_1 \; x_2 \; \dots x_d]^T$$

Let w_i be the classes where $i = 1, 2 \dots c$

$$P(x|w_i) = \prod_{j=1}^{d} p_{ij}^{x_j} (1 - p_{ij})^{(1 - x_j)} \quad \text{— (1)}$$

we know that

$$g_i(x) = P(w_i|x) - 1$$

Teacher's Signature ....

$g_i(x) = P(\omega_i | x)$

$g_i(x) = P(x|\omega_i) \cdot P(\omega_i)$

taking ln on both sides

$\ln g_i(x) = \ln P(x|\omega_i) + \ln P(\omega_i)$

from eq (i)

$g_i(x) = \ln P(x|\omega_i) + \ln P(\omega_i)$

$g_i(x) = \ln \left[\prod_{j=1}^{d} p_{ij}^{x_j} (1-p_{ij})^{1-x_j} \right] + \ln P(\omega_i)$

$$\boxed{g_i(x) = \sum_{j=1}^{d} \left[x_j \ln p_{ij} + (1-x_j) \ln (1-p_{ij}) + \ln P(\omega_i) \right]}$$

formulae for discrimiant of class ω_i

Given, for class 1

Prior $= P(\omega_1) = P(\omega_2) = 0.5$

$p_1 = \begin{bmatrix} 0.5 \\ 0.8 \end{bmatrix}$ $\quad p_{11} = 0.5$ $\quad p_{12} = 0.8$ $\quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

Therefore

$g_1(x) = \sum_{j=1}^{2} \left[x_j \ln p_{ij} + (1-x_j) \ln (1-p_{ij}) \right] + \ln 0.5$

$g_1(x) = x_1 \ln p_{11} + (1-x_1) \ln (1-p_{11}) + x_2 \ln p_{12}$
$+ (1-x_2) \ln (1-p_{12}) + \ln 0.5$

$= x_1 \ln (0.5) + (1-x_1) \ln (0.5) + x_2 \ln 0.8 + (1-x_2) \ln 0.2$
$+ \ln (0.5)$

$= (x_1 + 1 - x_1) \ln 0.5 + x \ln 0.8 + (1 - x_2) \ln 0.2 + \ln 0.5$

$= x_2 [\ln 0.8 - \ln 0.2] + \ln 0.2 + \ln 0.5 + \ln 0.5$

$= 1.386 x_2 - 2.995$

$\Rightarrow \boxed{g_1(x) = 1.386 x_2 - 2.995}$

For class 2

$$p_2 = \begin{bmatrix} 0.9 \\ 0.2 \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$p_{21} = 0.9 \quad \& \quad p_{22} \; 0.2$$

$$g_2(x) = \sum_{j=1}^{2} \left[x_j \ln p_{2j} + (1 - x_j) \ln (1 - p_{2j}) \right] + \ln(0.5)$$

$= x_1 \ln p_{21} + (1 - x_1) \ln (1 - p_{21}) + x_2 \ln p_{22} + (1 - x_2)$
$\ln (1 - p_{22}) + \ln 0.5$

$= x_1 \ln(0.9) + (1 - x_1) \ln(0.9) + x_2 \ln(0.2) + (1 - x_2)$
$\ln(0.8) + \ln(0.5)$

$= x_1 [\ln 0.9 - \ln 0.1] + x_2 [\ln 0.2 - \ln 0.8] +$
$\ln 0.1 + \ln 0.8 + \ln 0.5$

$= 2.197 x_1 - 1.386 x_2 - 3.218$

$$\boxed{g_2(x) = 2.194 x_1 - 1.386 x_2 - 3.218}$$

Teacher's Signature ....

3) a) Let us assume an arbitary matrix

$$X = \begin{bmatrix} 6 & 6 \\ 0 & 6 \end{bmatrix}$$

$$u = \begin{bmatrix} (6+6)/2 \\ (0+6)/2 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

computing centralized $(x) = x+c$

$$x-c = x-u = \begin{bmatrix} 6-6 & 6-6 \\ 0-3 & 6-3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 \\ -3 & 3 \end{bmatrix}$$

covariance matrix

$$S = \frac{1}{N}\left(x-c \cdot x-c^{t}\right)$$

$$= \frac{1}{2}\begin{bmatrix} 0 & 0 \\ -3 & 3 \end{bmatrix}\begin{bmatrix} 0 & -3 \\ 0 & 3 \end{bmatrix}$$

$$\frac{1}{2}\left\{\begin{bmatrix} 0 & 0 \\ 0 & 18 \end{bmatrix}\right\} = \begin{bmatrix} 0 & 0 \\ 0 & 9 \end{bmatrix}$$

$$S = \begin{bmatrix} 0 & 0 \\ 0 & 9 \end{bmatrix}$$

computing eignvalues & eignvector for above matrix

Let λ_1 & λ_2 be the eign values for eign values.

Teacher's Signature

$$|S - \lambda I| = 0$$

$$\begin{vmatrix} 0-\lambda & 0 \\ 0 & 9-\lambda \end{vmatrix} = 0$$

$\Rightarrow -\lambda(9-\lambda) = 0$

$\Rightarrow \lambda_1 = 0 \qquad \lambda_2 = 9$

Eign vectors $\vec{v}$

$$(A - \lambda I) \cdot \vec{v} = 0$$

for $\lambda_1 = 9$

$$(S - \lambda I)\ \vec{v} = 0$$

$$\vec{v} = \begin{bmatrix} a \\ b \end{bmatrix}$$

thus

$$\begin{bmatrix} -9 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0$$

$$-9a = 0$$

$$a = 0$$

$$b = b$$

$$\vec{v} = \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} = b \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\vec{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

for $\lambda_2 = 0$

$$(S - \lambda I)\ \vec{v} = 0$$

$$\vec{v} = \begin{bmatrix} c \\ d \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 \\ 0 & 9 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = 0$$

$$\begin{bmatrix} 0 \\ 9d \end{bmatrix} = 0$$

Teacher's Signature ....

$9d = 0$

$d = 0$

$c = c$

$$\vec{v} = \begin{bmatrix} c \\ d \end{bmatrix} = c\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Eignvalues = $9, 0$

Eign vectors = $\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

for $\lambda_1 = 9$ $\lambda_2 = 0$

$$U = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = [\vec{v_1}, \vec{v_2}]$$

eignvectors arranged acc to descending order of eignvalues.

$$U = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$Y = U^T . X - C$$

$$U^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad X - C = \begin{bmatrix} 0 & 0 \\ -3 & 3 \end{bmatrix}$$

$$Y = U^T . X - C$$
$$= \begin{bmatrix} -3 & 3 \\ 0 & 0 \end{bmatrix}$$

$$Y = \begin{bmatrix} -3 & 3 \\ 0 & 0 \end{bmatrix}$$

b) Let $A = UY + \text{mean}(X)$ $\quad B = X$

$A = UY + \text{mean}(X)$

$U = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $\quad Y = \begin{bmatrix} -3 & 3 \\ 0 & 0 \end{bmatrix}$ $\quad \text{mean}(X) = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$

$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -3 & 3 \\ 0 & 0 \end{bmatrix} + \text{mean}(X)$

$= \begin{bmatrix} 0 & 0 \\ -3 & 3 \end{bmatrix} + \begin{bmatrix} 6 \\ 3 \end{bmatrix}$

$A = \begin{bmatrix} 6 & 6 \\ 0 & 6 \end{bmatrix}$ $\quad B = \begin{bmatrix} 6 & 6 \\ 0 & 6 \end{bmatrix}$

finding MSE b/w A & B

$MSE = \frac{1}{4} \sum_{i=1}^{2} \sum_{j=1}^{2} (a_{ij} - b_{ij})^2$

$= \frac{1}{4} \sum_{i=1}^{2} \sum_{j=1}^{2} \left(\begin{bmatrix} 6 & 6 \\ 0 & 6 \end{bmatrix} - \begin{bmatrix} 6 & 6 \\ 0 & 6 \end{bmatrix} \right)$

$= \frac{1}{4} \sum_{i=1}^{2} \sum_{j=1}^{2} \left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \right)$

$= \frac{1}{4} \times 0$

$\boxed{MSE = 0}$