27-02-22

DATE
PAGE

Assignment 2

Q1 (b) MLE for multivariate Bernoulli distribution

$$P(x_j|\theta_j) = \theta_j^{x_j}(1-\theta_j)^{1-x_j}$$

$$\Rightarrow P(X|\theta_j) = \prod_{j=1}^{d} \theta_j^{x_j}(1-\theta_j)^{1-x_j} \quad \text{(for d-dimensions)}$$

N-iids of the distribution

$$Q(X) = \prod_{i=1}^{N}\prod_{j=1}^{d} \theta_j^{x_{ij}}(1-\theta_j)^{1-x_{ij}}$$

The log-likelihood function

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

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

Putting above to 0

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

Thus, for ~~th~~ j-th dimension

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

DATE
PAGE

(e) Taking discriminant for zero-one loss.
Here, the minimum error-rate classification is achieved by the function

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

i-th class

Here $P(\omega_1) = P(\omega_2)$

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

$$P(x_i|\omega_1) = \theta_i^{x_i}(1-\theta_i)^{1-x_i}$$

For d-dimensions

$$P(x|\omega_1) = \prod_{i=1}^{d} \theta_i^{x_i}(1-\theta_i)^{1-x_i}$$

Now, for j-th class

$$g_j(x) = \ln P(x|\omega_j)$$

$$\therefore P(x|\omega_j) = \prod_{i=1}^{d} \theta_{ij}^{x_i}(1-\theta_{ij})^{1-x_i}$$

$$\Rightarrow g_j(x) = \ln \prod_{i=1}^{d} \theta_{ij}^{x_i}(1-\theta_{ij})^{1-x_i}$$

$$\Rightarrow \boxed{g_j(x) = \sum_{i=1}^{d} x_i \ln \theta_{ij} + (1-x_i)\ln(1-\theta_{ij})}$$

DATE
PAGE

Q2 For Bernoulli

$$P(x_i|\theta_j) = \cancel{x_j}\ \theta_j^{x_i}(1-\theta_j)^{1-x_i}$$

(a) Now, for d-dimensions

$$P(X|\theta_j) = \prod_{j=1}^{d} \theta_j^{x_j}(1-\theta_j)^{1-x_j}$$

N-iids

$$Q(X) = \prod_{i=1}^{N}\prod_{j=1}^{d} \theta_j^{x_{ij}}(1-\theta_j)^{1-x_{ij}}$$

Log-likelihood when considering prior

$$F(\theta_j) = \ln(Q(X)\,P(\theta))$$

$$\Rightarrow F(\theta_j) = \ln\left\{\left[\prod_{i=1}^{N}\prod_{j=1}^{d} \theta_j^{x_{ij}}(1-\theta_j)^{1-x_{ij}}\right]P(\theta)\right\}$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{d}\left[x_{ij}\ln\theta_j + (1-x_{ij})\ln(1-\theta_j)\right] + \sum_{j=1}^{d}\ln\theta_j - \sum_{j=1}^{d}\theta_j$$

Now, $\dfrac{\partial F(\theta_j)}{\partial\theta_j} = 0$

$$\Rightarrow \sum_{i=1}^{N} x_{ij}\cdot\frac{1}{\theta_j} - \sum_{i=1}^{N}(1-x_{ij})\cdot\frac{1}{1-\theta_j} + \frac{1}{\theta_j} - 1 = 0$$

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

$$\Rightarrow \frac{X}{\theta_j} - \frac{(N-X)}{1-\theta_j} + \frac{1}{\theta_j} = 1$$

Ans. $\Rightarrow$ $\boxed{\theta_j^2 - (2+N)\theta_j + (X+1) = 0}$

(b) $X = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$

$d = 2 \; ; \quad N = 4$

θ_{MAP} is a 2×1 vector.

$$\theta_1^2 - (2+N)\theta_1 + \sum_{i=1}^{N} x_{i1} + 1 = 0$$

$$\Rightarrow \theta_1^2 - 6\theta_1 + 4 = 0$$

$$\Rightarrow \boxed{\theta_1 = 3 \pm \sqrt{5}}$$

Now, $\theta_2^2 - (2+N)\theta_2 + \sum_{i=1}^{N} x_{i2} + 1 = 0$

$$\Rightarrow \theta_2^2 - 6\theta_2 + 2 = 0$$

$$\Rightarrow \boxed{\theta_2 = 3 \pm \sqrt{7}}$$

$\because$ θ_i is probability, thus $\theta_i < 1$

$$\Rightarrow \theta_1 = 3 - \sqrt{5} \cong 0.764$$
$$\Rightarrow \theta_2 = 3 - \sqrt{7} \cong 0.354$$

Ans. $\Rightarrow \theta_{MAP} = \begin{bmatrix} 3-\sqrt{5} \\ 3-\sqrt{7} \end{bmatrix} \cong \begin{bmatrix} 0.764 \\ 0.354 \end{bmatrix}$

Q3. $X = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$; $\mu = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$

(a) $X_c = X - \mu = \begin{bmatrix} 2-3 & 4-3 \\ 6-7 & 8-7 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$

$S_{x_c} = \frac{1}{2}\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{N}\sum_{i=1}^{N} x_i x_i^T$

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

If λ are eigenvalues of S_{x_c}

$\det(S_{x_c} - \lambda I) = 0$

$\Rightarrow \det\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = 0$

$\Rightarrow \det\begin{pmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix} = 0$

$\Rightarrow (\lambda-1)^2 - 1^2 = 0$

$\Rightarrow (\lambda-1-1)(\lambda-1+1) = 0$

$\Rightarrow \lambda = 2 \quad ; \lambda = 0$

eigenvalues are $\boxed{2 \text{ and } 0}$

Let A be the eigenvector corresponding to $\lambda = 2$

$(S_{x_c} - \lambda I) A = 0$

$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = 0$

$\Rightarrow -a_1 + a_2 = 0 \quad \Rightarrow a_1 = a_2$

$$A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

B be the eigenvector corresponding to $\lambda = 0$.

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = 0$$

$\Rightarrow \quad b_1 + b_2 = 0 \quad \Rightarrow \quad b_1 = -b_2$

$\Rightarrow \quad B = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

Normalising A and B for PCA

$$\boxed{A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } B = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}}$$

$$\Rightarrow \boxed{U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}}$$

Taking 1 principal component : $p = 1$.

$\Rightarrow \quad Y = U_1^T . X_c$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 2 \end{bmatrix} 1/\sqrt{2}$$

$\Rightarrow \quad X_1' = U_1 . Y + \mu$

$$= \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} -2 & 2 \end{bmatrix} + \mu = \frac{1}{2} \begin{bmatrix} -2 & 2 \\ -2 & 2 \end{bmatrix} + \mu = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} + \mu$$

$$\Rightarrow \boxed{X_1' = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}}$$

(b) $MSE(X_1', X)$

$\because X_1' == X$

$\Rightarrow MSE = 0$ (as shown below)

$$\frac{(2-2)^2 + (4-4)^2 + (6-6)^2 + (8-8)^2}{4} = 0$$

Ans $\boxed{0}$

Now, taking two principal components : $p = 2$

$\Rightarrow Y = U^T X_c$

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

$\Rightarrow X_2' = U.Y + \mu$

$$= \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ +1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 3 \\ 7 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} -2 & 2 \\ -2 & +2 \end{bmatrix} + \begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$$

Again, $X_2' == X$

$\Rightarrow \boxed{MSE = 0}$