ASSIGNMENT-2

DEEPAM SARMAH
2020050

1.) Given $X = \begin{bmatrix} 1 & 3 \\ 4 & 7 \end{bmatrix}$; $X \in \mathbb{R}^{2\times 2}$; $d=2, n=2$.

Let $y_i = u_1^T x_i$ such that $u_1^T u_1 = 1$.
Now centralize x along mean i.e $X \leftarrow X - \mu$.
Now we want to maximise $\sigma_y^2 = u_1^T \Sigma u_1$ along u_1 (maximise energy) such that $u_1^T u_1 = 1$.
Using Lagrange Multipliers $\frac{\partial}{\partial u_1}\{u_1^T \Sigma u_1 - \lambda (u_1^T u_1 - 1)\} = 0$

$\Rightarrow 2\Sigma u_1 - 2\lambda u_1 = 0 \Rightarrow \Sigma u_1 = \lambda u_1$.

$\therefore$ If λ_1, λ_2 ($\lambda_1 \geq \lambda_2$) be the eigenvalues of Σ ~~then~~ corresponding to λ_1 & λ_2 are the eigen ~~values~~ vectors u_1 & u_2 respectively.

Now $\mu = \begin{bmatrix} \frac{1+3}{2} \\ \frac{4+7}{2} \end{bmatrix} = \begin{bmatrix} 2 \\ 5.5 \end{bmatrix}$.

Now $X \leftarrow X - \mu \Rightarrow X = \begin{bmatrix} 1-2 & 3-2 \\ 4-5.5 & 7-5.5 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -1.5 & 1.5 \end{bmatrix}$.

$\therefore \Sigma = \mathrm{Cov}(X) = \frac{1}{N} X X^T = \frac{1}{2} \begin{bmatrix} -1 & 1 \\ -1.5 & 1.5 \end{bmatrix} \begin{bmatrix} -1 & -1.5 \\ 1 & 1.5 \end{bmatrix}$

$\Rightarrow \Sigma = \frac{1}{2} \begin{bmatrix} 2 & 3 \\ 3 & 4.5 \end{bmatrix} = \begin{bmatrix} 1 & 1.5 \\ 1.5 & 2.25 \end{bmatrix}$.

Now $\det(\Sigma - \lambda I) = 0 \Rightarrow \begin{vmatrix} 1-\lambda & 1.5 \\ 1.5 & 2.25-\lambda \end{vmatrix} = 0 \Rightarrow \frac{\lambda}{4}(4\lambda - 13) = 0$

$\therefore \lambda_1 = \frac{13}{4}, \lambda_2 = 0$.

$\lambda_1 = 13/4$ we find $v_1 : \Sigma v_1 = \lambda v_1$.

$$\Rightarrow \begin{bmatrix} 1 & 1.5 \\ 1.5 & 2.25 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{13}{4} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

$$\Rightarrow \left. \begin{aligned} x_1 + 1.5x_2 &= 13/4\, x_1 \\ 1.5x_1 + 2.25x_2 &= 13/4\, x_2 \end{aligned} \right] \rightarrow x_1 = \frac{2}{3} x_2.$$

$$\therefore v_1 = \begin{bmatrix} 2/3 \\ 1 \end{bmatrix}.$$

$\lambda_2 = 0$ we find $v_2 : \Sigma v_2 = \lambda v_2$

$$\Rightarrow \begin{bmatrix} 1 & 1.5 \\ 1.5 & 2.25 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow \begin{aligned} 1x_1 + 1.5x_2 &= 0 \\ 1.5x_1 + 2.25x_2 &= 0 \end{aligned}$$

$$\Rightarrow x_1 + 3/2\, x_2 = 0 \quad \therefore v_2 = \begin{bmatrix} -3/2 \\ 1 \end{bmatrix}.$$

$\therefore$ PCA matrix is : $[v_1 \ v_2] = \begin{bmatrix} 2/3 & -3/2 \\ 1 & 1 \end{bmatrix}$

2.) Given $x_1, x_2 \in \mathbb{R}^{d \times N/2}$ so, $\mu_1, \mu_2 \in \mathbb{R}^{d \times 1}$.

Now projected samples be y_1 & y_2 such that

$y_1 = w^T x_1$, $y_2 = w^T x_2$

Now, given the projected samples follow a Bernoulli distribution and the samples are iid. Further, assume w is drawn from a multivariate Gaussian such that mean is 0 and covariance is identity matrix. Let's also assume that elements of w are statistically independent.

$\therefore$ $P(w) \propto e^{-w^T w/2}$.

Now MAP : $F(\theta) = \ln(P(y|\theta) \cdot P(w))$.

taking log likelihood, $F(\theta) = \ln(P(y|\theta)) + \ln P(w)$.

$\Rightarrow F(\theta) = \ln\left(\prod_{i=1}^{N} \theta^{y_i}(1-\theta)^{1-y_i}\right) + \ln\left(e^{-w^Tw/2}\right)$.

$\Rightarrow F(\theta) = \sum_{i=1}^{N}\left(y_i \ln\theta + (1-y_i)\ln(1-\theta)\right) + \ln\left(e^{-w^Tw/2}\right)$

$\Rightarrow F(\theta) = \sum_{i=1}^{N}\left(w^Tx_i \cdot \ln\theta + (1-w^Tx_i)\ln(1-\theta)\right) - w^Tw/2$

$\therefore$ we want to solve the following:

$$\max_{w} \sum_{i=1}^{N}\left(w^Tx_i \ln\theta + (1-w^Tx_i)\ln(1-\theta)\right) - \frac{w^Tw}{2} \quad \text{such that}$$

$$w^T\mu_1 = w^T\mu_2$$

By applying Lagrange Multipliers we get.

$$\frac{\partial}{\partial w}\cdot\left\{\sum_{i=1}^{N}\left(w^Tx_i \ln\theta + (1-w^Tx_i)\ln(1-\theta)\right\} - \frac{w^Tw}{2}\right. - \lambda \cdot \frac{\partial}{\partial w}\cdot(w^T\mu_1 - w^T\mu_2) = 0$$

$$\Rightarrow \sum_{i=1}^{N}\left(x_i \ln\theta - x_i\ln(1-\theta)\right) - w - \lambda(\mu_1 - \mu_2) = 0.$$

$$\Rightarrow w = \sum_{i=1}^{N}\left(x_i \ln\left(\frac{\theta}{1-\theta}\right)\right) - \lambda(\mu_1 - \mu_2) = 0.$$

Notice that $\mu_1 + \mu_2 = \frac{2}{N} \cdot \sum_{i=1}^{N} x_i$.

$\therefore \cancel{w = \sum_{i=1}^{N}(x_i}$ $\quad w = \frac{N}{2} \cdot \ln\left(\frac{\theta}{1-\theta}\right)(\mu_1 + \mu_2) - \lambda \cdot (\mu_1 - \mu_2)$

$$\boxed{w = \frac{N}{2}\ln\left(\frac{\theta}{1-\theta}\right)(\mu_1 + \mu_2) - \lambda(\mu_1 - \mu_2).}$$

3.) Given $P(x|w_2) \sim N(\mu, I)$

$P(x|w_3) \sim N(-\mu, I)$.

$$P(y|w_2; x) = \frac{1}{Z}\exp\left\{-0.5\,(y - u^T\mu)^T(y - u^T\mu)\right\}$$

$p(y|\omega_3; x) = \frac{1}{Z} \exp\{-0.5 (y+u^T\mu)^T (y+u^T\mu)\}$.

Now $S_w = S_2 + S_3 = (N-1)I + (N-1)I = 2(N-1)I$.

and $S = S_w + S_B \Rightarrow S_B = S - S_w = S - 2(N-1)I$ ①

$\therefore$ according to the question we want to :

$\max_u \ln p(y|\omega_2 : x)\, p(y|\omega_3 ; x)$ and also apply FDA conditions i.e $\max_u \frac{u^T S_B u}{u^T S_w u}$.

$\Rightarrow \cancel{\max_w \frac{\ln[p(y|\omega_2 : x)\, p(y|\omega_3 ; x)]\, u^T S_B u}{u^T S_w u}}$

$\Rightarrow \max_w \ln p(y|\omega_2 : x) \cdot p(y|\omega_3 : x) \cancel{*} \cancel{\frac{u^T S_B u}{u^T S_w u}}$

$$\Rightarrow \max_u \Big\{ -\frac{1}{2} [y^Ty - y^Tu^T\mu - u^T\mu y + u^T\mu\mu^T u + y^Ty + y^Tu^T\mu + u^T\mu y + u^T\mu\mu^T u] \cancel{+} \cancel{\frac{u^T S_B u}{u^T S_w u}} \Big\}.$$

$\therefore \max_u (-u^T\mu\mu^T u)$ & $\max_u \frac{u^T S_B u}{u^T S_w u}$.

$\therefore$ Applying Lagrange multipliers we get.

$$\frac{\partial}{\partial u} (-u^T\mu\mu^T u + u^T S_B u) - \lambda \frac{\partial}{\partial u} (u^T S_w u) = 0$$

$$\Rightarrow -2\mu\mu^T u + 2S_B u - \lambda (2 S_w u) = 0$$

$$\Rightarrow (S_B - \mu\mu^T) u = \lambda (S_w \cdot u)$$

we can see that u is a eigen vector of $S_w^{-1}(S_B - \mu\mu^T)$.

$\therefore u \propto S_W^{-1} \cdot [S_B - \mu\mu^T]$

$\Rightarrow u \propto S_W^{-1} [S - S_W - \mu\mu^T]$ ($\because$ from eq (1)).

$\Rightarrow u \propto S_W^{-1} [S - (2)(N-1) I \cdot - \mu\mu^T]$

$\Rightarrow u \propto \frac{I}{2(N-1)} [S - 2(N-1)I - \mu\mu^T]$

($\because S_W = 2(N-1)I \Rightarrow S_W^{-1} = \frac{I}{2(N-1)} \cdot$).