Date: 22/02/23

SML Assignment-2

Drishya Uniyal
MT21119

Theory.

B. 1. PCA weight matrix:-

$$\begin{vmatrix} 1 & 3 \\ 4 & 7 \end{vmatrix}$$

① Center the data by subtracting the mean from each data point.

$\mu = [(1+4)/2, (3+7)/2] = [2.5, 5]$

centered data matrix:

$X = [-1.5 \quad -2$
$\quad 1.5 \quad 2]$

Covariance matrix:

$\Sigma = X^T X/(n-1)$

\# of data points $n = 2$

$X^T X = [(-1.5)(1.5) + 1.5(1.5),$
$(-1.5)(-2) + 1.5(2);$
$(-2)(-1.5) + 2(1.5);$
$(-2)(-2) + 2(4)]$

$= \cancel{\begin{bmatrix} 3 & 3 \\ 8 & 8 \end{bmatrix}} \begin{bmatrix} 4.5 & -6 \\ -6 & 8 \end{bmatrix}$

Cov $\Sigma = \cancel{[3 \; 3, 3 \; 8]}/(2-1)$

$= \cancel{[3 \; 3, 3 \; 8]}$ = Cov-matrix $[[4.5 \; -6$
$-6 \; 8]]$

For Eigen Vectors & values of $\Sigma$

$\det(\Sigma - \lambda I) = 0$

Date: / / /

[[4.5 -6][[v1]][[λ* v1]]

[[-6 8]] @[v2]] = [λ * v2]]

~~det([3-λ 3; 3 8-λ]) = 0~~

~~$\lambda^2 - 11\lambda + 21 = 0$~~

~~$\lambda_1 = 2$~~

~~$\lambda_2 = 9$~~

~~Eigen Vector for $\lambda_1 = 2$~~

~~$(\Sigma - 2I)\, v_1 = 0$~~

~~$[1\ 3;\ 3\ 6]\, v_1 = 0$~~

~~$v_1 = [-3;\ 1]$~~

~~$\lambda_2 = 9$~~

~~$v_2 = [1;\ 3]$~~

~~$[-6\ 3;\ 3\ -1]\, v_2 = 0$~~

~~PCA $w = [-3\ 1;\ 1\ 3]$~~

direction.

$4.5 v_1 - 6 v_2 = \lambda v_1$  $-6 v_1 + 8 v_2 = \lambda v_2$

2. simplifying we get $\lambda_1 = 2$

$\lambda_2 = 16$

For $\lambda_1 = 2$

$v_1 = 12$

$v_2 = 5$

Eigen Vector [12, 5]

For $\lambda_2 = 16$  $(4.5 - 16) v_1 - 6 v_2 = 0$

$v_1 = -4$

$v_2 = 3$

Eigen vector [-4, 3]

Select Principal Comp.

Corresponding to largest eigen value [12 5] & [-4 3]

weight matrix $\begin{bmatrix} 12 & -4 \\ 5 & 3 \end{bmatrix}$

2. FDA weight vector $\omega$

$X1$ $X2$ data matrices

$c1$ $c2$ classes

# of samples in each class $N/2$ $N/2$

Follows Bernoulli dist $\theta$ iids

Constⁿ $w^T\mu_1 = w^T\mu_2$

$W = ?$

FDA → maximizes the between class scatter and min within class scatter achieved by finding a projection vector $w$ that maximizes Fischer

$$F(w) = (w^T\mu_1 - w^T\mu_2)^2 / w^T\Sigma w$$

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

$$= ((w^T\mu_1 + w^T\mu_2) - (w^T\mu_1 - w^T\mu_2)/2)^2 / w^T\Sigma w$$

$$= (w^T(\mu_1+\mu_2)/2)^2 / w^T\Sigma w$$

$$= (\mu_1+\mu_2)^T w w^T(\mu_1+\mu_2) / (w^T\Sigma w)$$

$$= w^T(\mu_1\mu^T + \mu_1\mu_2^T + \mu_2\mu_1^T + \mu_2\mu_2^T) w / w^T\Sigma w$$

Bernoulli

$$\Sigma = \theta(1-\theta)I$$

$$= (\mu_1\mu_1^T + \mu_1\mu_2^T + \mu_2\mu_1^T + \mu_2\mu_2^T) / \theta(1-\theta))$$

For $\omega$ that min. criterio

$$\frac{\partial F(\omega)}{\partial \omega} = 0$$

$$= 2(\mu_1\mu_1^T + \mu_1\mu_2^T + \mu_2\mu_1^T + \mu_2\mu_2^T) / (\theta(1-\theta)) -$$
$$2\theta(1-\theta) w (w^T(\mu_1\mu_1^T + \mu_1\mu_2^T + \mu_2\mu_1^T + \mu_2\mu_2^T) w)^1$$
$$-2(w^T(\theta((1-\theta)I) w)$$

Date: / /

Setting it equal to zero

$(\mu_1\mu_1^T + \mu_1\mu_2^T + \mu_2\mu_1^T + \mu_2\mu_2^T)w = \theta(1-\theta)Nw$

total samples $N1 + N2$

$w = (\theta(1-\theta)N)^{\wedge} - (\mu_1\mu_1^T + \mu_1\mu_2^T + \mu_2\mu_1^T + \mu_2\mu_2^T)^{\wedge} -1(\mu_1 - \mu_2)$

3. $P(x|\omega 2) \sim N(\mu_1 I)$
$P(x|\omega_3) \sim N(\mu_1 I)$

FDA → max b/w class variance &
min within " "

$S = S2 + S3$
Scatter ↓ ↓
$\omega 2$ $\omega 3$

$S = \frac{1}{N} \Sigma (x-\mu)(x-\mu)^T$

N ↓ No. of samples in class

$\mu$ ↓ mean

$S_b = (\mu_2 - \mu_3)(\mu_2 - \mu_3)^T$

means

FDA max ratio.
$J(u) = u^T S_b u / u^T S_w u$
↓
projection direction.

Max $J(u)$ equivalent to finding eigen vector of $Sw^{\wedge}1 - Sb$

Date: / / /

To incorporate Gaussian assumption we can compute the likelihood of the projections of y given the class conditions $p(y/\omega_2)$ & $p(y/\omega_3)$

$p(y/\omega_2) = (2\pi)^{-d/2} |Sw|^{-1/2} \exp(-0.5 (y-\mu_2)^T Sw^{-1} (y-\mu_2))$

$p(y/\omega_3) = (2\pi)^{-d/2} |Sw|^{-1/2} \exp(-0.5 (y-\mu_3)^T Sw^{-1} (y-\mu_3))$

Using $p(y/\omega_2; x)$ & $p(y/\omega_3; x)$
means $u^T\mu$ and $-u^T\mu$ ; covariance $I$

$p(y/\omega_2) = (2\pi)^{-d/2} |I|^{-1/2} \exp(-0.5 (y-u^T\mu)^T I^{-1} (y-u^T\mu))$

$p(y/\omega_3) = (2\pi)^{-d/2} |I|^{-1/2} \exp(-0.5 (y+\mu^T\mu)^T I^{-1} (y+\mu^T\mu))$

where $I \rightarrow 2\times 2$ matrix

max ~~likelihoods~~ log likelihoods

$\max_w [-0.5 (y-u^T\mu)^T I^{-1} (y-u^T\mu) - 0.5 (y+u^T\mu)^T I^{-1} (y+u^T\mu)]$

Expanding

$\max_w [-y^T I^{-1} y - 2\mu^T I^{-1} u^T y - u^T\mu^T I^{-1} \mu\mu]$

$w[-2\mu^T I^{-1} u^T y - u^T\mu^T I^{-1} \mu\mu]$

taking derivative w.r.t $u$ & $=0$

$-2\mu^T I^{-1} y + 2\mu^T I^{-1} \mu\mu = 0$

Date: / /

Now we have been given

$p(y/\omega_2) = 1/2\ \cancel{e}\ \exp(-0.5(y - u^T\mu_2)^T Sw^{\wedge}-1(y - u^T\mu_2))$

$p(y/\omega_3) = 1/2\ \exp(-0.5(y - u^T\mu_3)^T Sw^{\wedge}-1(y - u^T\mu_3))$

log-likelihood

$= +\ln(p(y|\omega_2; x)\,p(y|\omega_3; x)) = \ln(p(y|\omega_2) + \ln(p(y|\omega_3)$

$= -\ln 2 - 0.5(y - u^T(\mu_2 + \mu_3)/2)^T Sw^{\wedge}-1(y - u^T(\mu_2 + \mu_3)/2) - 0.5(\mu_2 - \mu_3)^T Sw^{\wedge}-1(\mu_2 - \mu_3)/4 - 0.5(u^T(\mu_2 - \mu_3))^T Sw^{\wedge}-1(\mu^T(\mu_2 - \mu_3)))/4$

$\mu_2 - \mu_3 = \cancel{\&}\ 2\mu_2 - (\mu_2 + \mu_3) - \mu_3 = \mu^T(\mu_2 - \mu_3)/2$

Max log likelihood; take derivative, equate to zero.

This gives $\boxed{u = Sw^{\wedge}-1(\mu_2 - \mu_3)}$