Data Mining

HW 2

(1) From the given information, we need to prove that

$$\|C_i\| \, \|x - \mu_i\|^2 = \sum_{x_j \in C_i} \|x_j - x\|^2 - \sum_{x_j \in C_i} \|x_j - \mu_i\|^2$$

where μ_i is the mean $\frac{\sum x_j}{n} = \frac{1}{C_i} \sum x_j$

n = size of cluster C_i

x = point in the dataset.

$$\sum_{x_j \in C_i} \|x_j - x\|^2 - \sum_{x_j \in C_i} \|x_j - \mu_i\|^2.$$

$$= \sum_{x_j \in C_i} \left(\cancel{x_j^2} + x^2 - 2 \cdot x_j \cdot x - \cancel{x_j^2} + \mu_i^2 - 2 \cdot x_j \mu_i \right)$$

$$= \sum_{x_j \in C_i} \left(x^2 - \mu_i^2 - 2 x_j (x - \mu_i) \right)$$

$$= \sum_{x_j \in C_i} x^2 - \sum_{x_j \in C_i} \mu_i^2 - \sum_{x_j \in C_i} 2 x_j (x - \mu_i)$$

$$= C_i x^2 - C_i \mu_i^2 - 2(\mu_i C_i)(x - \mu_i)$$

$$= C_i x^2 - C_i \mu_i^2 - 2 \mu_i C_i x + 2 C_i \mu_i^2.$$

$$= C_i x^2 + C_i \mu_i^2 - 2 \cdot \mu_i C_i x .$$

$$= C_i x^2 - 2 \cdot C_i \, x \cdot \mu_i + C_i \mu_i^2$$

$$= \|C_i\| \, (x^2 - 2 \cdot x \cdot \mu_i + \mu_i)^2$$

$$(\because a^2 - 2ab + b^2 = (a-b)^2)$$

$$= \|C_i\| \cdot \|x - \mu_i\|^2$$

$= L.H.S$

$=$ Hence proved.

(2) Q. L.D.A We ignored $w^T \cdot w = 1$.

Use lagrange multiplier

we've

$$J(w) = \frac{w^T B w}{w^T S w}$$

$B =$ Scatter matrix (Betweencluster)

$S = S_1 + S_2$.

Lagrange multiplier, with the $w^T w = 1$.

$$L(w, \lambda) = J(w) - \lambda(w^T . w - 1)$$

Taking partial derivation w.r.t w. & equal to 0

$$\frac{\partial L}{\partial w} = \frac{\partial}{\partial w}(J(w)) - \frac{\partial}{\partial w}\lambda(w^T . w - 1) . = 0$$

$$\Rightarrow \frac{\partial}{\partial w}\left[\frac{w^T . B . w}{w^T . S . w}\right] - \frac{\partial}{\partial w}\{\lambda . (w^T . w - 1)\} = 0$$

here, w is a unit-vector

$w^T . w \Rightarrow$ Column vector, $w \, [w_1 \; w_2 \; \cdots \; w_n]$.

$$w^T = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \Rightarrow w^T . w = w_1^2 + w_2^2 + \cdots w_n^2$$

$$\therefore \frac{\partial}{\partial w} w^T . w = 2 . (w_1 + w_2 \cdots w_n)$$

$$= 2 . w . n . \approx 2 . W .$$

$$\therefore \frac{\partial L}{\partial w} \Rightarrow 0$$

$$\Rightarrow \frac{2 . B . w . (w^T . Sw) - S . w (w^T . B . w)}{(w^T . S . w)^2} - 2 . \lambda w = 0$$

$$\frac{\partial}{\partial w} J(w) = \frac{\partial}{\partial w}\left[\frac{w^T . B . w}{w^T . S . w}\right]$$

$$= \frac{\frac{\partial}{\partial w}(w^T . B . w) \times w^T S . w - w^T B . w \frac{\partial}{\partial w}(w^T S$$

$$(w^T . S . w)^2$$

$$\frac{\partial}{\partial w}(w^T . B . w) = (B + B^T) w \cong 2 . B . w$$

$$\frac{\partial}{\partial w}(w^T . S . w) = (S + S^T) w \cong 2 . S . w$$

$$\therefore \frac{\partial}{\partial w}(J(w)) = \frac{w^T . S . w (B + B^T) w - w^T B w (S + S^T}{(w^T . S . w)^2}$$

$$\frac{\partial(L)}{\partial w} = \frac{\partial}{\partial w}(J(w)) - \frac{\partial}{\partial w} \lambda (w^T w - 1)$$

$$\Rightarrow \frac{(B + B^T) w (w^T S . w) - (w^T B . w)(S + S^T) w}{(w^T . S . w)^2} - 2 . \lambda w = 0$$

Applying Numerical method optimization,

$$\frac{2Bw(w^T \cdot S \cdot w) - (w^T B w)\, 2 \cdot S \cdot w}{(w^T \cdot S \cdot w)^2} - 2\lambda w = 0$$

$$2Bw \cdot (w^T S \cdot w) - (w^T B \cdot w)(2 \cdot S \cdot w) = 2 \cdot \lambda w (w^T S w)^2$$

$$\frac{2 \cdot B \cdot w (w^T \cdot S \cdot w) - (w^T B \cdot w)(2 \cdot S \cdot w)}{2 \cdot \lambda \cdot w} = (w^T \cdot S \cdot w)^2$$

$$\frac{\cancel{2}\left(B \cdot w (w^T \cdot S \cdot w) - (w^T B \cdot w) S \cdot w\right)}{(w^T \cdot S \cdot w)^2} = \cancel{2} \cdot \lambda \cdot w$$

$$\therefore \quad \lambda \cdot w = \frac{B \cdot w \cdot (w^T \cdot S \cdot w) - S \cdot w (w^T B \cdot w)}{(w^T \cdot S \cdot w)^2}$$

$$w = \frac{B \cdot w (w^T \cdot S \cdot w) - S \cdot w (w^T B w)}{(\lambda \cdot (w^T \cdot S \cdot w)^2)}$$

$$\frac{\partial L}{\partial \lambda} = 0$$

$$\frac{\partial}{\partial L}\left(J(w) - \lambda (w^T w - 1)\right) = 0.$$

$$w^T \cdot w - 1 = 0 \Rightarrow w^T \cdot w = 1.$$

Substituting value of w in here,

$$\left[\frac{B \cdot w\,(w^T S \cdot w) - (w^T B\, w)\, S\, w}{\lambda (w^T S \cdot w)^2}\right]^T \left[\frac{B \cdot w\,(w^T \cdot S \cdot w) - (w^T \cdot B \cdot w)\, S}{\lambda \cdot (w^T \cdot S \cdot w)^2}\right]$$

$$= 1$$

(3) given,

$$D = \begin{array}{ccc} x_1 & x_2 & \text{class} \end{array}$$

$$D = \begin{bmatrix} 2 & 3 & 0 \\ 3 & 3 & 0 \\ 3 & 4 & 0 \\ 5 & 8 & 0 \\ 7 & 7 & 0 \\ 5 & 4 & 1 \\ 6 & 5 & 1 \\ 7 & 4 & 1 \\ 7 & 5 & 1 \\ 8 & 2 & 1 \\ 9 & 4 & 1 \end{bmatrix}$$

<u>For class 0</u> :- Mean of $x_1 = \frac{2+3+3+5+7}{5} = 4$

Mean of $x_2 = \cancel{4+5+} \frac{3+3+4+8+7}{5} = 5$

<u>Class 1</u> mean $x_1 = \frac{5+6+7+7+8+9}{6} \cong 6.83$

mean of $x_2 = \frac{4+5+4+5+2+4}{6} = 4.$

Scatter Matrix:-

<u>For class 0</u> :- $S_W^{(0)} = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$

$$S_W^{(1)} = \begin{bmatrix} 1.47 & 0.17 \\ 0.17 & 1.83 \end{bmatrix}$$

Scanned with CamScanner

$$S_W = S_W^{(0)} + S_W^{(1)} = \begin{bmatrix} 8.47 & 1.17 \\ 1.17 & 5.83 \end{bmatrix}$$

B/W Scatter matrix B : Overall $= \begin{bmatrix} 5.5 \\ 4.5 \end{bmatrix}$

$$B = \begin{bmatrix} -13.5 & -5.5 \\ -5.5 & -2.25 \end{bmatrix}$$

Eigenvectors & Eigenvalues

eigenvalues $= S^{-1}B = \begin{bmatrix} -5.29 & -2.14 \\ -12.39 & -5.10 \end{bmatrix}$

$$\lambda_1, \lambda_2 = 0.00078, -9.398$$

eigenvectors $x_1 = \begin{bmatrix} 0.924 \\ 0.383 \end{bmatrix}$ $x_2 = \begin{bmatrix} -0.383 \\ 0.924 \end{bmatrix}$

Discriminant direction

largest eigenvalue $= \begin{bmatrix} 0.924 \\ 0.383 \end{bmatrix}$

$$W = \frac{\lambda_1(1)}{\lambda_2(0)} = \frac{0.383}{0.924} = 0.415$$

B = (mean of class 1) - mx
(mean of class 0)

$b = -247.$

$y = 0.415x - 2.47.$