a. class 1: $\mu_{11} = \frac{1}{n_1}\sum x_i = \frac{1}{4}(0.8+0.9+1.2+1.1) = 1$

$\mu_{12} = \frac{1}{n_1}\sum y_i = \frac{1}{4}(1.2+1.4+1.4+1.5) = 1.375$

$$\mu_1 = \begin{bmatrix}\mu_{11}\\ \mu_{12}\end{bmatrix} = \begin{bmatrix}1\\ 1.375\end{bmatrix}$$

$\sigma_{11} = \frac{1}{3}\left((1-0.8)^2+(1-0.9)^2+(1-1.2)^2+(1-1.1)^2\right)$

$= \frac{1}{3}\times(0.2^2+0.1^2+0.2^2+0.1^2)$

$= 0.033$

$\sigma_{22} = \frac{1}{3}\times\left((1.375-1.2)^2+(1.375-1.4)^2+(1.375-1.4)^2+(1.375-1.5)^2\right)$

$= \frac{1}{3}\times(0.175^2+0.025^2+0.025^2+0.125^2)$

$= 0.158$

$\sigma_{12} = \frac{1}{3}\times\left((1-0.8)\times(1.375-1.2)+(1-0.9)\times(1.375-1.4)+(1-1.2)\times(1.375-1.4)+(1-1.1)\times(1.375-1.5)\right)$

$= \frac{1}{3}\times 0.05$

$= 0.017$

$\sigma_{21} = \sigma_{12} = 0.017$

in conclusion, the mean of class 1 is $\mu_1 = \begin{bmatrix}1\\ 1.375\end{bmatrix}$

the covariance is $\Sigma_1 = \begin{bmatrix}0.033 & 0.017\\ 0.017 & 0.158\end{bmatrix}$

for class 2.

$$\Sigma_2 = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \frac{1}{n_2 - 1} \sum_{k=1}^{4} (\vec{x}_k - \vec{\mu}_2)(\vec{x}_k - \vec{\mu}_2)^T$$

$$\mu_2 = \begin{bmatrix} \frac{1}{n_2} \Sigma X_{21} \\ \frac{1}{n_2} \Sigma X_{22} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{4}(0.8 + 0.6 + 0.65 + 0.75) \\ \frac{1}{4}(1.1 + 1.0 + 1.1 + 0.9) \end{bmatrix}$$

$$= \begin{bmatrix} 0.7 \\ 1.025 \end{bmatrix}$$

$$= \frac{1}{3} \times \Sigma \left(\begin{bmatrix} 0.8 - 0.7 \\ 1.1 - 1.025 \end{bmatrix} [0.8 - 0.7, 1.1 - 1.025] + \begin{bmatrix} 0.6 - 0.7 \\ 1 - 1.025 \end{bmatrix} [0.6 - 0.7, 1 - 1.025] + \begin{bmatrix} 0.65 - 0.7 \\ 1.1 - 1.025 \end{bmatrix} [0.65 - 0.7, 1.1 - 1.025] + \begin{bmatrix} 0.75 - 0.7 \\ 0.9 - 1.025 \end{bmatrix} [0.75 - 0.7, 0.9 - 1.025] \right.$$

$$= \frac{1}{3} \times \Sigma \left(\begin{bmatrix} 0.1^2 & 0.1 \times 0.075 \\ 0.1 \times 0.075 & 0.075^2 \end{bmatrix} + \begin{bmatrix} 0.1^2 & 0.1 \times 0.025 \\ 0.1 \times 0.025 & 0.025^2 \end{bmatrix} + \begin{bmatrix} 0.05^2 & -0.05 \cdot 0.075 \\ -0.05 \cdot 0.075 & 0.075^2 \end{bmatrix} + \begin{bmatrix} 0.05^2 & -0.05 \cdot 0.125 \\ -0.05 \cdot 0.125 & 0.125^2 \end{bmatrix} \right)$$

$$= \frac{1}{3} \times \Sigma \begin{bmatrix} 0.025 & 0 \\ 0 & 0.0275 \end{bmatrix}$$

$$= \begin{bmatrix} 0.008 & 0 \\ 0 & 0.009 \end{bmatrix}$$

To summarize, $\mu_2 = \begin{bmatrix} 0.7 \\ 1.025 \end{bmatrix}$ $\Sigma_2 = \begin{bmatrix} 0.008 & 0 \\ 0 & 0.009 \end{bmatrix}$

1) C

Mahalanobis distance.

$$d_M = \sqrt{(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})} \quad \text{where } \vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \ \vec{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

$$d_{M1} = \sqrt{\left(\begin{bmatrix} 0.85 \\ 1.15 \end{bmatrix} - \begin{bmatrix} 1 \\ 1.375 \end{bmatrix}\right)^T \cdot \begin{bmatrix} 0.033 & 0.017 \\ 0.017 & 0.016 \end{bmatrix} \left(\begin{bmatrix} 0.85 \\ 1.15 \end{bmatrix} - \begin{bmatrix} 1 \\ 1.375 \end{bmatrix}\right)}$$

$$= \sqrt{\begin{bmatrix} -0.15, & -0.225 \end{bmatrix} \begin{bmatrix} 0.033 & 0.017 \\ 0.017 & 0.016 \end{bmatrix} \begin{bmatrix} -0.15 \\ -0.225 \end{bmatrix}}$$

$$= \sqrt{\begin{bmatrix} -0.15 \cdot 0.033 - 0.225 \times 0.017, & -0.15 \cdot 0.017 - 0.225 \times 0.016 \end{bmatrix} \begin{bmatrix} -0.15 \\ -0.225 \end{bmatrix}}$$

$$= \sqrt{\begin{matrix} 0.15^2 \cdot 0.033 + 0.225 \times 0.017 \times 0.15 + 0.15 \times 0.225 \cdot 0.033 + 0.225^2 \cdot 0.017 \\ + 0.15^2 \cdot 0.017 + 0.225 \cdot 0.016 \cdot 0.15 + 0.225 \cdot 0.016 \cdot 0.15 + 0.225^2 \cdot 0.016 \end{matrix}}$$

$$d_{M2} = \sqrt{\left(\begin{bmatrix} 0.85 \\ 1.15 \end{bmatrix} - \begin{bmatrix} 0.7 \\ 1.025 \end{bmatrix}\right)^T \begin{bmatrix} 0.008 & 0 \\ 0 & 0.009 \end{bmatrix} \left(\begin{bmatrix} 0.85 \\ 1.15 \end{bmatrix} - \begin{bmatrix} 0.7 \\ 1.025 \end{bmatrix}\right)}$$

$$= \sqrt{\begin{bmatrix} 0.15 & 0.125 \end{bmatrix} \begin{bmatrix} 0.008 & 0 \\ 0 & 0.009 \end{bmatrix} \begin{bmatrix} 0.15 \\ 0.125 \end{bmatrix}}$$

$$= \sqrt{\begin{bmatrix} 0.15 \times 0.008 & 0.125 \times 0.009 \end{bmatrix} \begin{bmatrix} 0.15 \\ 0.125 \end{bmatrix}}$$

$$= \sqrt{0.15^2 \cdot 0.008 + 0.125^2 \cdot 0.009}$$

Euclidean distance

$$d_{e_1} = \sqrt{(\vec{x}-\vec{\mu})^T(\vec{x}-\vec{\mu})}$$

$$= \sqrt{\left(\begin{bmatrix}0.85\\1.15\end{bmatrix}-\begin{bmatrix}1\\1.375\end{bmatrix}\right)^T\left(\begin{bmatrix}0.85\\1.15\end{bmatrix}-\begin{bmatrix}1\\1.375\end{bmatrix}\right)}$$

$$= \sqrt{\begin{bmatrix}-0.15 & -0.225\end{bmatrix}\begin{bmatrix}-0.15\\-0.225\end{bmatrix}}$$

$$= \sqrt{0.15^2+0.225^2}$$

$$d_{e_2} = \sqrt{\left(\begin{bmatrix}0.85\\1.15\end{bmatrix}-\begin{bmatrix}0.7\\1.025\end{bmatrix}\right)^T\left(\begin{bmatrix}0.85\\1.15\end{bmatrix}-\begin{bmatrix}0.7\\1.025\end{bmatrix}\right)}$$

$$= \sqrt{\begin{bmatrix}0.15 & 0.125\end{bmatrix}\begin{bmatrix}0.15\\0.125\end{bmatrix}}$$

$$= \sqrt{0.15^2+0.125^2}$$

1) c Mahalanobis distance:

assume the test sample is $\begin{bmatrix} a \\ b \end{bmatrix}$, class i mean $= \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$

class i covariance $= \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$

where $\sigma_{12} = \sigma_{21}$

the Mahalanobis distance is

$$d_m = \sqrt{(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})}$$

$$= \sqrt{\left(\begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right)^T \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}^{-1} \left(\begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right)}$$

$$= \sqrt{\begin{bmatrix} a-\mu_1 & b-\mu_2 \end{bmatrix} \cdot \frac{1}{\sigma_{11}\sigma_{22}-\sigma_{21}\cdot\sigma_{12}} \cdot \begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{21} & \sigma_{11} \end{bmatrix} \cdot \begin{bmatrix} a-\mu_1 \\ b-\mu_2 \end{bmatrix}}$$

$$= \sqrt{\frac{1}{\sigma_{11}\sigma_{22}-\sigma_{21}\cdot\sigma_{12}} \begin{bmatrix} (a-\mu_1)\cdot\sigma_{22} - (b-\mu_2)\cdot\sigma_{21}, & -(a-\mu_1)\sigma_{12} + (b-\mu_2)\cdot\sigma_{11} \end{bmatrix} \cdot \begin{bmatrix} a-\mu_1 \\ b-\mu_2 \end{bmatrix}}$$

$$= \sqrt{\frac{1}{\sigma_{11}\sigma_{22}-\sigma_{21}\cdot\sigma_{12}}} \cdot \sqrt{\begin{aligned} &(a-\mu_1)^2\sigma_{22} - (b-\mu_2)(a-\mu_1)\sigma_{21} + (a-\mu_1)\sigma_{22}\cdot(b-\mu_2) \\ &- (b-\mu_2)^2\sigma_{21} + \\ &- (a-\mu_1)^2\sigma_{12} + (b-\mu_2)(a-\mu_1)\sigma_{11} - (a-\mu_1)(b-\mu_2)\sigma_{12} \\ &+ (b-\mu_2)^2\sigma_{11} \end{aligned}}$$

$\xrightarrow{\sigma_{12}=\sigma_{21}}$

$$\sqrt{\frac{1}{\sigma_{11}\sigma_{22}-\sigma_{21}\cdot\sigma_{12}}} \cdot \sqrt{\begin{aligned} &(b-\mu_2)(a+b-\mu_1-\mu_2)\sigma_{11} - \left[(a-\mu_1)^2+(b-\mu_2)^2+2(a-\mu_1)(b-\mu_2)\right]\sigma_{21} \\ &+ (a-\mu_1)(a+b-\mu_1-\mu_2)\sigma_{22} \end{aligned}}$$

$\xrightarrow{\sigma_{12}=\sigma_{21}}$

$$\sqrt{\frac{1}{\sigma_{11}\sigma_{22}-\sigma_{21}^2}} \cdot \sqrt{\begin{aligned} &(b-\mu_2)(a+b-\mu_1-\mu_2)\sigma_{11} - (a+b-\mu_1-\mu_2)^2\sigma_{21} \\ &+ (a-\mu_1)(a+b-\mu_1-\mu_2)\sigma_{22} \end{aligned}}$$

Euclidean distance

$$d_e = \sqrt{(\vec{x}-\vec{\mu})^T(\vec{x}-\vec{\mu})}$$

$$= \sqrt{\left(\begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right)^T \left(\begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right)}$$

$$= \sqrt{\begin{bmatrix} a-\mu_1 & b-\mu_2 \end{bmatrix} \begin{bmatrix} a-\mu_1 \\ b-\mu_2 \end{bmatrix}}$$

$$= \sqrt{(a-\mu_1)^2+(b-\mu_2)^2}$$

① use dmin method

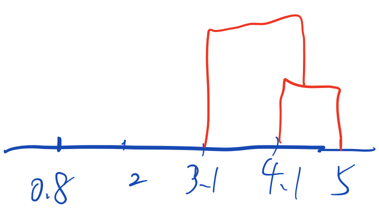
value	0.8	2	3.1	4.1	5
id	0	1	2	3	4

step 1.

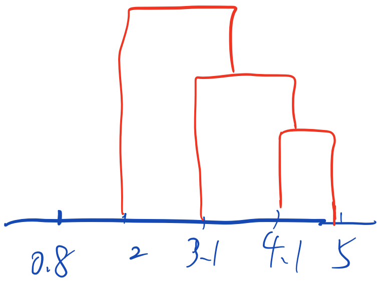
	0	1	2	3	4
0	∅	1.2	2.3	3.3	4.2
1	∞	∅	1.1	2.1	3
2	X	X	∅	1	1.9
3	X	X	X	∅	0.9
4	X	X	X	X	∅

step 2.

	0	1	2	34
0	∅	1.2	2.3	min[(0,3), (0,4)] = 3.3
1	∅	∅	1.1	min[(1,3),(1,4)] = 2.1
2	∞	X	∅	min[(2,3),(2,4)] = 1
34	X	X	X	∅

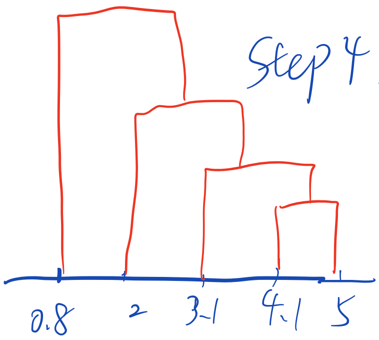


step 3.



	0	1	234
0	×	1.2	min[(0,2),(0,3),(0,4)] = 2.3
1	×	×	min[(1,2),(1,3)(1,4)] = 1.1
234	×	×	×

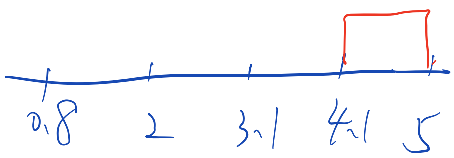
Step 4.



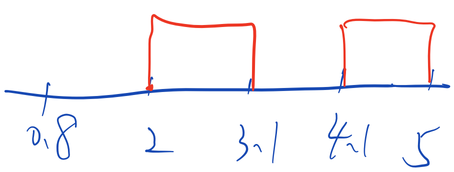
	0	1234
0	×	min[(0,1)(0,2)(0,3)(0,4)] = 1.2
1234	×	×

② use dmax method: use the maximum distance between clusters elements while use minimum distance to incorporate new element.

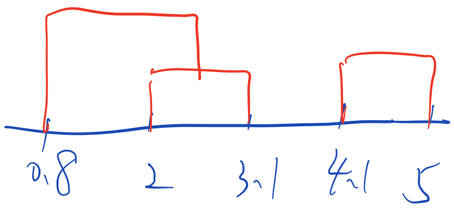
value	0.8	2	3.1	4.1	5
id	0	1	2	3	4



	0	1	2	3	4
0	X	1.2	2.3	3.3	4.2
1	X	X	1.1	2.1	3
2	X	X	X	1	1.9
3	X	X	X	X	0.9
4	X	X	X	X	X

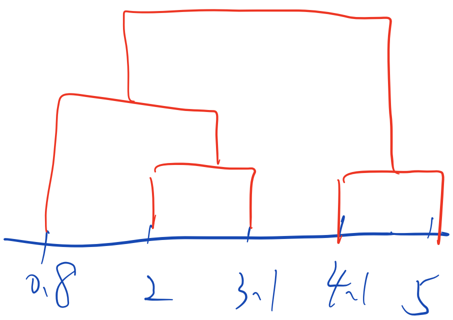


	0	1	2	34
0	X	1.2	2.3	max[(0,3)(0,4)]=4.2
1	X	X	1.1	max[(1,3)(1,4)]=3
2	X	X	X	max[(2,3)(2,4)]=1.9
34	X	X	X	X



max((0,1) (0,2)) = 2.3

	0	12	34
0	∅	2.3	4.2
12	X	∅	3
34	X	X	∅



	012	34
012	∅	4.2
34	*	*

4 use MLE to derive the equation for mean and variance assuming 1-D Gussian .

$$p(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

likelihood $\rightarrow$ $L(\mu,\sigma) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \cdot \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2\right)$

log likelihood $\rightarrow$ $\ell(\mu,\sigma) = \log(L(\mu,\sigma))$

$$= n\cdot\log 1 - n\cdot\log\sqrt{2\pi}\,\sigma - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2$$

$$= -\frac{n}{2}\log 2\pi - n\log\sigma - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2$$

$$\frac{\partial \ell(\mu,\sigma)}{\partial\mu} = -\frac{1}{2}\sum_{i=1}^{n}\cdot 2\left(\frac{x_i-\mu}{\sigma}\right)\cdot\left(-\frac{1}{\sigma}\right) = \sum_{i=1}^{n}\frac{x_i-\mu}{\sigma^2} = 0 \qquad (1)$$

$$\frac{\partial \ell(\mu,\sigma)}{\partial\sigma} = -\frac{n}{\sigma} - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\cdot\frac{-2}{\sigma^3} \qquad (2)$$

$$= -\frac{n}{\sigma} + \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^3} = 0$$

From equation (1) we can get $\sum_{i=1}^{n} x_i - \mu = 0$

$$\mu = \frac{\sum x_i}{n}$$

solve equation (2), we get $n\sigma^2 = \sum_{i=1}^{n}(x_i-\mu)^2$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2}$$

where μ is solved by equation (1)