Name: Chaoyang Zhu UIN: 930001840

Problem 1: K-means

Initial:

1）Cluster assignment

$C_1=\{x_1, x_2, x_4\}$

$C_2=\{x_3, x_5\}$

2）Compute means

$$\mu_1=(\frac{5}{3}, \frac{2}{3})$$

$$\mu_2=(\frac{13}{4}, 1)$$

Iteration 1:

$$\|X_1-\mu_1\|_2=\sqrt{(0-\frac{5}{3})^2+(2-\frac{2}{3})^2}=\frac{\sqrt{41}}{3}<\|X_1-\mu_2\|_2=\sqrt{(0-\frac{13}{4})^2+(2-1)^2}=\frac{\sqrt{185}}{4}$$

$$\|X_2-\mu_1\|_2=\sqrt{(0-\frac{5}{3})^2+(0-\frac{2}{3})^2}=\frac{\sqrt{29}}{3}<\|X_2-\mu_2\|_2=\sqrt{(0-\frac{13}{4})^2+(0-1)^2}=\frac{\sqrt{185}}{4}$$

$$\|X_3-\mu_1\|_2=\sqrt{(\frac{3}{2}-\frac{5}{3})^2+(0-\frac{2}{3})^2}=\frac{\sqrt{17}}{6}<\|X_3-\mu_2\|_2=\sqrt{(\frac{3}{2}-\frac{13}{4})^2+(0-1)^2}=\frac{\sqrt{65}}{4}$$

$$\|X_4-\mu_1\|_2=\sqrt{(5-\frac{5}{3})^2+(0-\frac{2}{3})^2}=\frac{\sqrt{104}}{3}>\|X_4-\mu_2\|_2=\sqrt{(5-\frac{13}{4})^2+(0-1)^2}=\frac{\sqrt{65}}{4}$$

$$\|X_5-\mu_1\|_2=\sqrt{(5-\frac{5}{3})^2+(2-\frac{2}{3})^2}=\frac{\sqrt{116}}{3}>\|X_5-\mu_2\|_2=\sqrt{(5-\frac{13}{4})^2+(2-1)^2}=\frac{\sqrt{65}}{4}$$

1）Cluster assignment

$C_1=\{x_1, x_2, x_3\}$

$C_2=\{x_4, x_5\}$

2）Compute means

$$\mu_1=(\frac{1}{2}, \frac{2}{3})$$

$$\mu_2=(5, 1)$$

Iteration 2:

$$\left\|x_1 - \mu_1\right\|_2 = \sqrt{(0 - \frac{1}{2})^2 + (2 - \frac{2}{3})^2} = \frac{\sqrt{73}}{6} < \left\|x_1 - \mu_2\right\|_2 = \sqrt{(0-5)^2 + (2-1)^2} = \sqrt{26}$$

$$\left\|x_2 - \mu_1\right\|_2 = \sqrt{(0 - \frac{1}{2})^2 + (0 - \frac{2}{3})^2} = \frac{5}{6} < \left\|x_2 - \mu_2\right\|_2 = \sqrt{(0-5)^2 + (0-1)^2} = \sqrt{26}$$

$$\left\|x_3 - \mu_1\right\|_2 = \sqrt{(\frac{3}{2} - \frac{1}{2})^2 + (0 - \frac{2}{3})^2} = \frac{\sqrt{13}}{3} < \left\|x_3 - \mu_2\right\|_2 = \sqrt{(\frac{3}{2}-5)^2 + (0-1)^2} = \frac{\sqrt{53}}{2}$$

$$\left\|x_4 - \mu_1\right\|_2 = \sqrt{(5 - \frac{1}{2})^2 + (0 - \frac{2}{3})^2} = \frac{\sqrt{745}}{6} > \left\|x_4 - \mu_2\right\|_2 = \sqrt{(5-5)^2 + (0-1)^2} = 1$$

$$\left\|x_5 - \mu_1\right\|_2 = \sqrt{(5 - \frac{1}{2})^2 + (2 - \frac{2}{3})^2} = \frac{\sqrt{793}}{6} > \left\|x_5 - \mu_2\right\|_2 = \sqrt{(5-5)^2 + (2-1)^2} = 1$$

Cluster assignment

$C_1=\{x_1, x_2, x_3\}$

$C_2=\{x_4, x_5\}$

The same as iteration 1

So the process has converged.

The final clusters are given as **$C_1=\{x_1, x_2, x_3\}$** and **$C_2=\{x_4, x_5\}$**.

Problem 2: Gaussian Mixture Models

(A)

$$w_{ij} = \frac{f(x_j \mid \mu_i, \Sigma_i) \cdot P(C_i)}{\sum_{a=1}^{k} f(x_j \mid \mu_a, \Sigma_a) \cdot P(C_a)}$$

Since k = 3, $\Sigma_1 = \Sigma_2 = \Sigma_3 = \{ \{1,0\}, \{0,1\} \} = \Sigma$ and $P(C_1) = P(C_2) = P(C_3) = 1/3$.

$$w_{ij} = \frac{f(x_j \mid \mu_i, \Sigma) \cdot P(C_i)}{f(x_j \mid \mu_1, \Sigma) \cdot P(C_1) + f(x_j \mid \mu_2, \Sigma) \cdot P(C_2) + f(x_j \mid \mu_3, \Sigma) \cdot P(C_3)}$$

$$= \frac{f(x_j \mid \mu_i, \Sigma)}{f(x_j \mid \mu_1, \Sigma) + f(x_j \mid \mu_2, \Sigma) + f(x_j \mid \mu_3, \Sigma)}$$

According to multivariate normal distribution,

$$f(x \mid \mu_i, \Sigma_i) = \frac{1}{(2\pi)^{\frac{d}{2}} \left|\Sigma_i\right|^{\frac{1}{2}}} \exp\{- \frac{(x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i)}{2}\}$$

In the problem, d=2 and $\Sigma_1 = \Sigma_2 = \Sigma_3 = \{ \{1,0\}, \{0,1\} \} = \Sigma$

So, $f(x \mid \mu_i, \Sigma_i) = \dfrac{1}{2\pi} \exp\{- \dfrac{(x - \mu_i)^T (x - \mu_i)}{2}\}$

Hence,

$$w_{ij} = \frac{\frac{1}{2\pi}\exp\{-\frac{(x_j-\mu_i)^T(x_j-\mu_i)}{2}\}}{\frac{1}{2\pi}\exp\{-\frac{(x_j-\mu_1)^T(x_j-\mu_1)}{2}\} + \frac{1}{2\pi}\exp\{-\frac{(x_j-\mu_2)^T(x_j-\mu_2)}{2}\} + \frac{1}{2\pi}\exp\{-\frac{(x_j-\mu_3)^T(x_j-\mu_3)}{2}\}}$$

$$= \frac{\exp\{-\frac{(x_j-\mu_i)^T(x_j-\mu_i)}{2}\}}{\exp\{-\frac{(x_j-\mu_1)^T(x_j-\mu_1)}{2}\} + \exp\{-\frac{(x_j-\mu_2)^T(x_j-\mu_2)}{2}\} + \exp\{-\frac{(x_j-\mu_3)^T(x_j-\mu_3)}{2}\}}$$

In Matlab,

```
>> D=[0.5,4.5;2.2,1.5;3.9,3.5;2.1,1.9;0.5,3.2;0.8,4.3;2.7,1.1;2.5,3.5;2.8,3.9;0.1,4.1]
D =
    0.5000    4.5000
    2.2000    1.5000
    3.9000    3.5000
    2.1000    1.9000
    0.5000    3.2000
    0.8000    4.3000
    2.7000    1.1000
    2.5000    3.5000
    2.8000    3.9000
    0.1000    4.1000
```

Since initial mean $\mu_1 = (0.5, 4.5)^T$, $\mu_2 = (2.2, 1.6)^T$ and $\mu_1 = (3, 3.5)^T$

```
>> CenteredData1=D-repmat([0.5,4.5],10,1)
CenteredData1 =
         0         0
    1.7000   -3.0000
    3.4000   -1.0000
    1.6000   -2.6000
         0   -1.3000
    0.3000   -0.2000
    2.2000   -3.4000
    2.0000   -1.0000
    2.3000   -0.6000
   -0.4000   -0.4000
>> CenteredData2=D-repmat([2.2,1.6],10,1)
CenteredData2 =
```

```
   -1.7000    2.9000
         0   -0.1000
    1.7000    1.9000
   -0.1000    0.3000
   -1.7000    1.6000
   -1.4000    2.7000
    0.5000   -0.5000
    0.3000    1.9000
    0.6000    2.3000
   -2.1000    2.5000
>> CenteredData3=D-repmat([3,3.5],10,1)
CenteredData3 =
   -2.5000    1.0000
   -0.8000   -2.0000
    0.9000         0
   -0.9000   -1.6000
   -2.5000   -0.3000
   -2.2000    0.8000
   -0.3000   -2.4000
   -0.5000         0
   -0.2000    0.4000
   -2.9000    0.6000

>> CenteredData1Square=CenteredData1.^2
CenteredData1Square =
         0         0
    2.8900    9.0000
   11.5600    1.0000
    2.5600    6.7600
         0    1.6900
    0.0900    0.0400
    4.8400   11.5600
    4.0000    1.0000
    5.2900    0.3600
    0.1600    0.1600
>> CenteredData1SquareSum=CenteredData1Square(:,1)+CenteredData1Square(:,2)
CenteredData1SquareSum =
         0
   11.8900
   12.5600
    9.3200
```

```
    1.6900
    0.1300
   16.4000
    5.0000
    5.6500
    0.3200
>> CenteredData2Square=CenteredData2.^2
CenteredData2Square =
    2.8900    8.4100
         0    0.0100
    2.8900    3.6100
    0.0100    0.0900
    2.8900    2.5600
    1.9600    7.2900
    0.2500    0.2500
    0.0900    3.6100
    0.3600    5.2900
    4.4100    6.2500
>>   CenteredData2SquareSum=CenteredData2Square(:,1)+CenteredData2Square(:,2)
CenteredData2SquareSum =
   11.3000
    0.0100
    6.5000
    0.1000
    5.4500
    9.2500
    0.5000
    3.7000
    5.6500
   10.6600
>> CenteredData3Square=CenteredData3.^2
CenteredData3Square =
    6.2500    1.0000
    0.6400    4.0000
    0.8100         0
    0.8100    2.5600
    6.2500    0.0900
    4.8400    0.6400
    0.0900    5.7600
    0.2500         0
    0.0400    0.1600
```

```
    8.4100    0.3600
>> CenteredData3SquareSum=CenteredData3Square(:,1)+CenteredData3Square(:,2)
CenteredData3SquareSum =
    7.2500
    4.6400
    0.8100
    3.3700
    6.3400
    5.4800
    5.8500
    0.2500
    0.2000
    8.7700
>> e1=CenteredData1SquareSum;
>> e2=CenteredData2SquareSum;
>> e3=CenteredData3SquareSum;
>> w=zeros(3,10);
for j=1:10
w(1,j)=exp(-e1(j)/2)/(exp(-e1(j)/2)+exp(-e2(j)/2)+exp(-e3(j)/2));
w(2,j)=exp(-e2(j)/2)/(exp(-e1(j)/2)+exp(-e2(j)/2)+exp(-e3(j)/2));
w(3,j)=exp(-e3(j)/2)/(exp(-e1(j)/2)+exp(-e2(j)/2)+exp(-e3(j)/2));
end
>> w
w =
    0.9707    0.0024    0.0026    0.0083    0.7998    0.9265    0.0003    0.0732    0.0579    0.9801
    0.0034    0.9079    0.0548    0.8299    0.1220    0.0097    0.9352    0.1402    0.0579    0.0056
    0.0259    0.0897    0.9426    0.1618    0.0782    0.0638    0.0644    0.7867    0.8841    0.0143

>> w(1,:)*D
ans =
    2.1034    15.4407
>> TotalWeight1=0;
>> for j=1:10
TotalWeight1=TotalWeight1+w(1,j);
end
>> TotalWeight1
TotalWeight1 =
    3.8218
```

So, $\mu_1 = (\frac{2.1034}{3.8218}, \frac{15.4407}{3.8218})^T = (0.5503, 4.0402)^T$

```
>> w(2,:)*D
ans =
   7.0629    5.3463
>> TotalWeight2=0;
>> for j=1:10
TotalWeight2=TotalWeight2+w(2,j);
end
>> TotalWeight2
TotalWeight2 =
   3.0667
```

So, $\mu_2 = (\dfrac{7.0629}{3.0667}, \dfrac{5.3463}{3.0667})^T = (2.3031, 1.7433)^T$

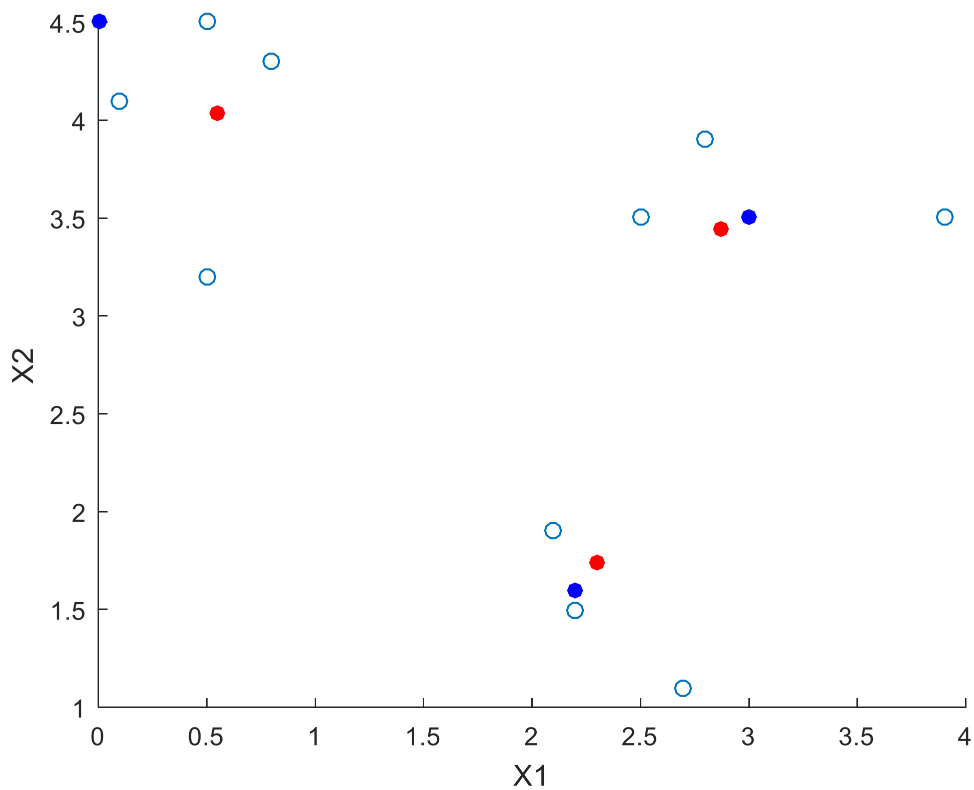
```
>> w(3,:)*D
ans =
   8.9337    10.7130
>> TotalWeight3=0;
>> for j=1:10
TotalWeight3=TotalWeight3+w(3,j);
end
>> TotalWeight3
TotalWeight3 =
   3.1115
```

So, $\mu_3 = (\dfrac{8.9337}{3.1115}, \dfrac{10.7130}{3.1115})^T = (2.8712, 3.4430)^T$

Hence, **μ1 = (0.5503, 4.0402)** $^{\mathbf{T}}$, **μ2 = (2.3031, 1.7433)** $^{\mathbf{T}}$ and **μ3 = (2.8712, 3.4430)** $^{\mathbf{T}}$.

(B)

```
>> X=D(:,1);
>> Y=D(:,2);
>> scatter(X,Y)
>> hold on;
>> xlabel('X1');
>> ylabel('X2');
>> scatter(0,4.5,'b','filled');
>> scatter(2.2,1.6,'b','filled');
>> scatter(3,3.5,'b','filled');
>> scatter(0.5503,4.0402,'r','filled');
>> scatter(2.3031,1.7433,'r','filled');
>> scatter(2.8712,3.4430,'r','filled');
```



Iterated means are more representative of clusters than initial means.

(C)

$$P^1(C_1) = \frac{\Sigma_{j=1}^{10} W_{1j}}{10} = \frac{TotalWeight1}{10} = \frac{3.8218}{10} = \mathbf{0.38}$$

$$P^1(C_2) = \frac{\Sigma_{j=1}^{10} W_{2j}}{10} = \frac{TotalWeight2}{10} = \frac{3.0667}{10} = \mathbf{0.31}$$

$$P^1(C_3) = \frac{\Sigma_{j=1}^{10} W_{3j}}{10} = \frac{TotalWeight3}{10} = \frac{3.1115}{10} = \mathbf{0.31}$$

(D)

$$\Sigma_i^1 = \frac{\Sigma_{j=1}^{n} W_{ij}(x_j - \mu_i)(x_j - \mu_i)^T}{\Sigma_{j=1}^{n} W_{ij}}$$

In Matlab,

```
>> CCenteredData1=D-repmat([0.5503,4.0402],10,1)
CCenteredData1 =
```

```
   -0.0503    0.4598
    1.6497   -2.5402
    3.3497   -0.5402
    1.5497   -2.1402
   -0.0503   -0.8402
    0.2497    0.2598
    2.1497   -2.9402
    1.9497   -0.5402
    2.2497   -0.1402
   -0.4503    0.0598
>> re_cov1 = zeros(2);
>> for j=1:10
re_cov1(1,1)=re_cov1(1,1)+w(1,j)*(CCenteredData1(j,1))^2;
re_cov1(1,2)=re_cov1(1,2)+w(1,j)*CCenteredData1(j,1)*CCenteredData1(j,2);
re_cov1(2,1)=re_cov1(1,2);
re_cov1(2,2)=re_cov1(2,2)+w(1,j)*(CCenteredData1(j,2))^2;
end
>> re_cov1=re_cov1/3.8218
re_cov1 =
    0.2329   -0.0247
   -0.0247    0.2395
```

So, $\boldsymbol{\Sigma_1^1} = \begin{bmatrix} \mathbf{0.2329} & \mathbf{-0.0247} \\ \mathbf{-0.0247} & \mathbf{0.2395} \end{bmatrix}$

```
>> CCenteredData2=D-repmat([2.3031,1.7433],10,1)
CCenteredData2 =
   -1.8031    2.7567
   -0.1031   -0.2433
    1.5969    1.7567
   -0.2031    0.1567
   -1.8031    1.4567
   -1.5031    2.5567
    0.3969   -0.6433
    0.1969    1.7567
    0.4969    2.1567
   -2.2031    2.3567
>> re_cov2 = zeros(2);
>> for j=1:10
re_cov2(1,1)=re_cov2(1,1)+w(2,j)*(CCenteredData2(j,1))^2;
re_cov2(1,2)=re_cov2(1,2)+w(2,j)*CCenteredData2(j,1)*CCenteredData2(j,2);
```

```
re_cov2(2,1)=re_cov2(1,2);
re_cov2(2,2)=re_cov2(2,2)+w(2,j)*(CCenteredData2(j,2))^2;
end
>> re_cov2=re_cov2/3.0667
re_cov2 =
    0.2633   -0.1245
   -0.1245    0.5581
```

So, $\boldsymbol{\Sigma}_2^1 = \begin{bmatrix} \mathbf{0.2633} & \mathbf{-0.1245} \\ \mathbf{-0.1245} & \mathbf{0.5581} \end{bmatrix}$

```
>> CCenteredData3=D-repmat([2.8712,3.4430],10,1)
CCenteredData3 =
   -2.3712    1.0570
   -0.6712   -1.9430
    1.0288    0.0570
   -0.7712   -1.5430
   -2.3712   -0.2430
   -2.0712    0.8570
   -0.1712   -2.3430
   -0.3712    0.0570
   -0.0712    0.4570
   -2.7712    0.6570
>> re_cov3 = zeros(2);
>> for j=1:10
re_cov3(1,1)=re_cov3(1,1)+w(3,j)*(CCenteredData3(j,1))^2;
re_cov3(1,2)=re_cov3(1,2)+w(3,j)*CCenteredData3(j,1)*CCenteredData3(j,2);
re_cov3(2,1)=re_cov3(1,2);
re_cov3(2,2)=re_cov3(2,2)+w(3,j)*(CCenteredData3(j,2))^2;
end
>> re_cov3=re_cov3/3.1115
re_cov3 =
    0.7129    0.0598
    0.0598    0.4353
```

So, $\boldsymbol{\Sigma}_3^1 = \begin{bmatrix} \mathbf{0.7129} & \mathbf{0.0598} \\ \mathbf{0.0598} & \mathbf{0.4353} \end{bmatrix}$