Question 1

The input data consists of the twenty 2D points $(i, i), (i, i+1)$ where $i = 1, 2, \ldots 10$

1. The mean of the input data is $\mu = (5.5, 6)$

   The new centered points are now of the form $(j, j-0.5), (j, j+0.5)$ where $j = -4.5, -3.5, \ldots, 3.5, 4.5$

2. The matrix X after centering is as follows:

$$X = \begin{bmatrix} -4.5 & -5 \\ -4.5 & -4 \\ \vdots & \vdots \\ 3.5 & 5 \\ 4.5 & 5 \\ 4.5 & 6 \end{bmatrix}$$

$$\text{i.e. X is the below matrix:}$$

$$X = \begin{bmatrix} j & j-0.5 \\ j & j+0.5 \end{bmatrix}_{20\times 2} \text{ where } j = -4.5, -3.5..., 3.5, 4.5$$

   In order to get the Sample Covariance Matrix, we calculate $S = \frac{1}{19}X^TX$:

$$\begin{aligned} X^TX &= \begin{bmatrix} j & j \\ j-0.5 & j+0.5 \end{bmatrix}_{2\times 20} \begin{bmatrix} j & j-0.5 \\ j & j+0.5 \end{bmatrix}_{20\times 2} \text{ where } j = -4.5, -3.5..., 3.5, 4.5 \\ &= \begin{bmatrix} \Sigma 2j^2 & \Sigma 2j^2 \\ \Sigma 2j^2 & \Sigma(2j^2 + 0.5) \end{bmatrix}_{2\times 2} \text{ where } j = -4.5, -3.5..., 3.5, 4.5 \\ &= \begin{bmatrix} 165 & 165 \\ 165 & 170 \end{bmatrix}_{2\times 2} \end{aligned}$$

$$S = \frac{1}{19}X^TX = \begin{bmatrix} 8.68421 & 8.68421 \\ 8.68421 & 8.94737 \end{bmatrix}_{2\times 2}$$

3. On performing Eigen decomposition of S, the two principal components obtained are:

$$v_1 = \begin{bmatrix} -0.702 \\ -0.712 \end{bmatrix}_{2\times 1} \text{ with } \lambda_1 = 17.501$$

$$\text{and}$$

$$v_2 = \begin{bmatrix} -0.712 \\ -0.702 \end{bmatrix}_{2\times 1} \text{ with } \lambda_2 = 0.131$$

4. Projecting each centered point onto v_1, we have

$$\begin{bmatrix} j & j-0.5 \end{bmatrix} \begin{bmatrix} -0.702 \\ -0.712 \end{bmatrix} \text{ and } \begin{bmatrix} j & j+0.5 \end{bmatrix} \begin{bmatrix} -0.702 \\ -0.712 \end{bmatrix}$$

$$\text{where } j = -4.5, -3.5..., 3.5, 4.5$$

Hence the one dimensional projection of these points(maximal spread) is of the form:

$$\begin{bmatrix} -1.414j + 0.356 \\ -1.414j - 0.356 \end{bmatrix}_{20\times 1} \text{ where } j = -4.5, -3.5..., 3.5, 4.5$$

Projecting each centered point onto v_2, we have

$$\begin{bmatrix} j & j-0.5 \end{bmatrix} \begin{bmatrix} -0.712 \\ 0.702 \end{bmatrix} \text{ and } \begin{bmatrix} j & j+0.5 \end{bmatrix} \begin{bmatrix} -0.712 \\ 0.702 \end{bmatrix}$$
$$\text{where } j = -4.5, -3.5..., 3.5, 4.5$$

which is of the form

$$\begin{bmatrix} -0.01j - 0.351 \\ -0.01j + 0.351 \end{bmatrix}_{20\times 1} \text{ where } j = -4.5, -3.5..., 3.5, 4.5$$

Hence, the 2D projection of the twenty points after PCA results in the new twenty points as given below:

$$\begin{bmatrix} -1.414j + 0.356 & -0.01j - 0.351 \\ -1.414j - 0.356 & -0.01j + 0.351 \end{bmatrix}_{20\times 2} \text{ where } j = -4.5, -3.5..., 3.5, 4.5$$

Question 2

The Knowledge Base KB is $p \implies (q \implies r)$

Converting to CNF,

$$KB : p \implies (\neg q \vee r) \text{.....Using implication elimination}$$
$$\neg p \vee (\neg q \vee r) \text{.....Using implication elimination}$$
$$(\neg p \vee \neg q \vee r) \text{.....Associativity principle of } \vee$$

The query β is $(p \wedge q) \implies (q \implies r)$

Converting to CNF,

$$\beta : \neg(p \wedge q) \vee (q \implies r) \text{.....Using implication elimination}$$
$$(\neg p \vee \neg q) \vee (\neg q \vee r) \text{.....Using implication elimination and DeMorgan principle}$$
$$\neg p \vee (\neg q \vee \neg q) \vee r \text{.....Using Associativity principle of } \vee$$
$$(\neg p \vee \neg q \vee r) \text{.....since a value ORed with itself gives the value itself}$$

It can be observed that the Knowledge Base and query give the same results for any values of p,q and r. As a result, every interpretation for which KB results in true, β will also result in true and hence KB entails β.

To show the steps more formally,

$$KB \wedge \neg\beta :$$
$$(\neg p \vee \neg q \vee r) \wedge \neg(\neg p \vee \neg q \vee r)$$
$$(\neg p \vee \neg q \vee r) \wedge (p \wedge q \wedge \neg r) \text{.....Using DeMorgan's Law}$$

Cancelling out the negated terms, we get empty. Hence the query is proved.

Question 3

Let the cities be represented by M,S,B,V,W,Mo for Madison, Seattle, Boston, Vancouver, Winnipeg and Montreal.

So there are initially 6 clusters: M,S,B,V,W,Mo.

1. The required parameters at each iteration are as follows:

a) Iteration 1

(1) The closest pair of clusters: V and S. Let VS be the new cluster formed.

(2) The distance between them as defined by complete linkage:

$$\sqrt{(-123+122)^2+(49-48)^2}=\sqrt{2}=1.414$$

(3) All clusters at the end of the iteration: VS, M, B, W, Mo

b) Iteration 2

(1) The closest pair of clusters: B and Mo. Let BMo be the new cluster formed.

(2) The distance between them as defined by complete linkage:

$$\sqrt{(-74+71)^2+(46-42)^2}=\sqrt{25}=5$$

(3) All clusters at the end of the iteration: VS, M, BMo, W

c) Iteration 3

(1) The closest pair of clusters: W and M. Let WM be the new cluster formed.

(2) The distance between them as defined by complete linkage:

$$\sqrt{(-97+89)^2+(50-43)^2}=\sqrt{113}=10.63$$

(3) All clusters at the end of the iteration: VS, BMo, WM

d) Iteration 4

(1) The closest pair of clusters: WM and BMo. Let WMBMo be the new cluster formed.

(2) The distance between them as defined by complete linkage: The furthest points between the two clusters are W and B.

$$\sqrt{(-97+71)^2+(50-42)^2}=\sqrt{26^2+8^2}=27.2$$

(3) All clusters at the end of the iteration: VS, WMBMo

Hence finally we arrive at two clusters: VS and WMBMo.

2. The required parameters at each iteration are as follows:

   a) Iteration 1

   (1) The closest pair of clusters: M and B. Let MB be the new cluster formed.

   (2) The distance between them as defined by complete linkage:

$$\sqrt{(-89+71)^2+(43-42)^2}=\sqrt{325}=18.03$$

   (3) All clusters at the end of the iteration: V, S, MB, W, Mo

   b) Iteration 2

   (1) The closest pair of clusters: W and Mo. Let WMo be the new cluster formed.

   (2) The distance between them as defined by complete linkage:

$$\sqrt{(-97+74)^2+(50-46)^2} = \sqrt{545} = 23.345$$

(3) All clusters at the end of the iteration: V, S, MB, WMo

c) Iteration 3

(1) The closest pair of clusters: V and WMo. Let VWMo be the new cluster formed.

(2) The distance between them as defined by complete linkage: The furthest two points between the two clusters are V and Mo.

$$\sqrt{(-123+74)^2+(49-46)^2} = \sqrt{2410} = 49.092$$

(3) All clusters at the end of the iteration: S, MB, VWMo

d) Iteration 4

(1) The closest pair of clusters: S and MB. Let SMB be the new cluster formed.

(2) The distance between them as defined by complete linkage: The furthest points between the two clusters are S and B.

$$\sqrt{(-122+71)^2+(48-42)^2} = \sqrt{2637} = 51.352$$

(3) All clusters at the end of the iteration: SMB, VWMo

Hence finally we arrive at two clusters: SMB and VWMo.

Question 4

1. Initially, $c_1 = 0, c_2 = 9$. The initial energy before clustering = 40. The required parameters at each iteration of the K-means clustering are as follows:

   a) Iteration 1

   (1) The cluster assignments are as follows:

   Cluster 1: x_4, x_5, x_6

   Cluster 2: x_1, x_2, x_3

   i.e. $y_1 = c_2, y_2 = c_2, y_3 = c_2, y_4 = c_1, y_5 = c_1, y_6 = c_1$

   (2) Updated clusters at the end of the iteration:

   $c_1 = \frac{2+3+4}{3} = 3$

   $c_2 = \frac{6+8+10}{3} = 8$

   (3) Energy at the end of the iteration:

$$\begin{gathered}(2-3)^2+(3-3)^2+(4-3)^2+(6-8)^2+(8-8)^2+(10-8)^2\\ = 1^2+0^2+1^2+2^2+0^2+2^2\\ = 10\end{gathered}$$

b) Iteration 2

The cluster centers are not updated, and hence we get same values as above.

Since the cluster centers have stopped moving, we stop K-means clustering at this point.

2. Initially, $c_1 = 8, c_2 = 9$. The initial energy before clustering = 82. The required parameters at each iteration of the K-means clustering are as follows:

   a) Iteration 1

   (1) The cluster assignments are as follows:

   Cluster 1: x_2, x_3, x_4, x_5, x_6

   Cluster 2: x_1

   i.e. $y_1 = c_2, y_2 = c_1, y_3 = c_1, y_4 = c_1, y_5 = c_1, y_6 = c_1$

   (2) Updated clusters at the end of the iteration:

   $c_1 = \frac{2+3+4+6+8}{5} = 4.6$

   $c_2 = \frac{10}{1} = 10$

   (3) Energy at the end of the iteration:

$$
\begin{gathered}
(2-4.6)^2 + (3-4.6)^2 + (4-4.6)^2 + (6-4.6)^2 + (8-10)^2 + (10-10)^2 \\
= 2.6^2 + 1.6^2 + 0.6^2 + 1.4^2 + 2^2 + 0^2 \\
= 15.64
\end{gathered}
$$

b) Iteration 2

(1) The cluster assignments are as follows:

Cluster 1: x_3, x_4, x_5, x_6

Cluster 2: x_1, x_2

i.e. $y_1 = c_2, y_2 = c_2, y_3 = c_1, y_4 = c_1, y_5 = c_1, y_6 = c_1$

(2) Updated clusters at the end of the iteration:

$c_1 = \frac{2+3+4+6}{4} = 3.75$

$c_2 = \frac{8+10}{2} = 9$

(3) Energy at the end of the iteration:

$$
\begin{gathered}
(2-3.75)^2 + (3-3.75)^2 + (4-3.75)^2 + (6-3.75)^2 + (8-9)^2 + (10-9)^2 \\
= 1.75^2 + 0.75^2 + 0.25^2 + 2.25^2 + 1^2 + 1^2 \\
= 10.75
\end{gathered}
$$

c) Iteration 3

The cluster centers are not updated, and hence we get same values as above.

Since the cluster centers have stopped moving, we stop K-means clustering at this point.

3. Clearly the first k-means solution seems to be a better choice. The second one required one extra step to complete execution as compared to the first and moreover it also ended up with a higher

distortion(energy) value. Hence, the better choice of picking the starting points for the cluster centers is to ensure they're further apart.