Q1-1: Given that you are using K-means clustering algorithm to obtain 3 clusters from 7 data points in 2-dim. In the first iteration, clusters C1, C2 and C3 are assigned data points as below.

C1: {(2,2),(4,4),(6,6)}, C2: {(0,4),(4,0)}, C3: {(5,5),(9,9)}

What will be the cluster centroids at the start of second iteration?

1. C1: (4,4), C2: (2,2), C3: (7,7)
2. C1: (6,6), C2: (4,4), C3: (9,9)
3. C1: (2,2), C2: (0,0), C3: (5,5)
4. C1: (2,6), C2: (0,4), C3: (5,9)

Q1-1: Given that you are using K-means clustering algorithm to obtain 3 clusters from 7 data points in 2-dim. In the first iteration, clusters C1, C2 and C3 are assigned data points as below.

C1: {(2,2),(4,4),(6,6)}, C2: {(0,4),(4,0)}, C3: {(5,5),(9,9)}

What will be the cluster centroids at the start of second iteration?

1. C1: (4,4), C2: (2,2), C3: (7,7) ⬅
2. C1: (6,6), C2: (4,4), C3: (9,9)
3. C1: (2,2), C2: (0,0), C3: (5,5)
4. C1: (2,6), C2: (0,4), C3: (5,9)

Q1-2: Consider the K-means algorithm with K = 3. After current iteration, we have 3 centers C1: (0,1), C2: (2,1), C3: (-1,2).

Which cluster assignment is possible for the points A: (1,1) and B: (-1,1) respectively? Assume ties are broken arbitrarily.

(i) C1, C1
(ii) C2, C3
(iii) C1, C3

1. Only (i)
2. Only (ii) and (iii)
3. Only (i) and (iii)
4. All of them

Q1-2: Consider the K-means algorithm with K = 3. After current iteration, we have 3 centers C1: (0,1), C2: (2,1), C3: (-1,2).

Which cluster assignment is possible for the points A: (1,1) and B: (-1,1) respectively? Assume ties are broken arbitrarily.

(i) C1, C1
(ii) C2, C3
(iii) C1, C3

1. Only (i)
2. Only (ii) and (iii)
3. Only (i) and (iii)
4. All of them 

Squared Euclidean distance between A and centers: 1, 1, 5

For B: 1, 9, 1

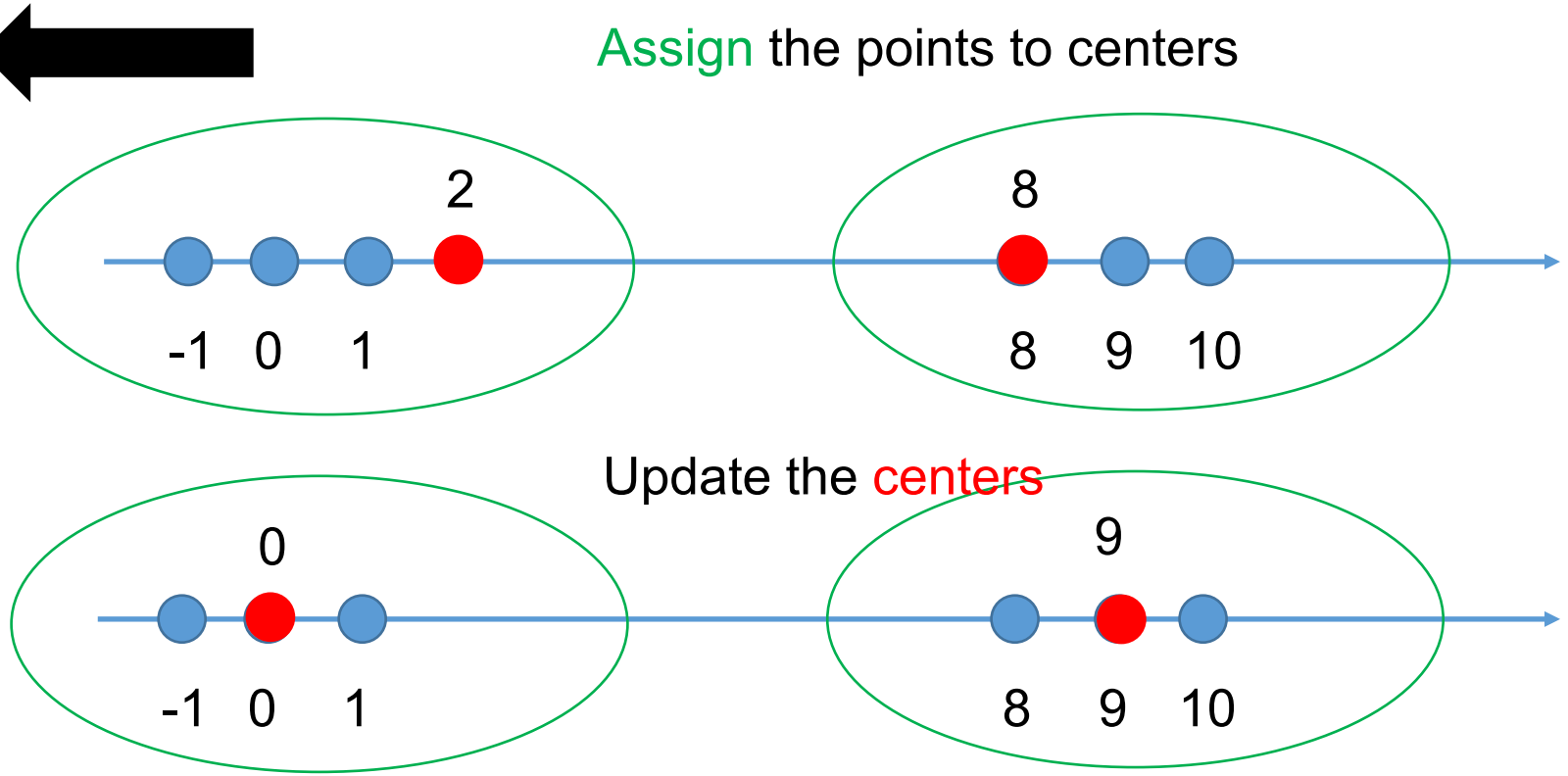
So A can be assigned to C1 and C2, B can be to C1 and C3

Q1-3: Given the following points in 1D: x1 = -1, x2 = 0, x3 = 1, x4 = 8, x5 = 9, x6 = 10, what are the locations of cluster centers at convergence assuming K=2? Assume we start with cluster centers c1 = 2 and c2 = 8.

1. c1 = 2, c2 = 8
2. c1 = 0, c2 = 9
3. c1 = -1, c2 = 10
4. c1 = 0, c2 = 0

Q1-3: Given the following points in 1D: x1 = -1, x2 = 0, x3 = 1, x4 = 8, x5 = 9, x6 = 10, what are the locations of cluster centers at convergence assuming K=2? Assume we start with cluster centers c1 = 2 and c2 = 8.

1. c1 = 2, c2 = 8
2. c1 = 0, c2 = 9
3. c1 = -1, c2 = 10
4. c1 = 0, c2 = 0



Q2-1: Consider the K-means algorithm from the slides. Which step changes cluster centers to minimize distortion?

1. Step 1
2. Step 2

Q2-1: Consider the K-means algorithm from the slides. Which step changes cluster centers to minimize distortion?

1. Step 1
2. Step 2 

Q2-2: Consider the K-means algorithm from the slides. Which step assigns each x to its closest cluster center y(x) to minimize the distortion?

1. Step 1
2. Step 2

Q2-2: Consider the K-means algorithm from the slides. Which step assigns each x to its closest cluster center y(x) to minimize the distortion?

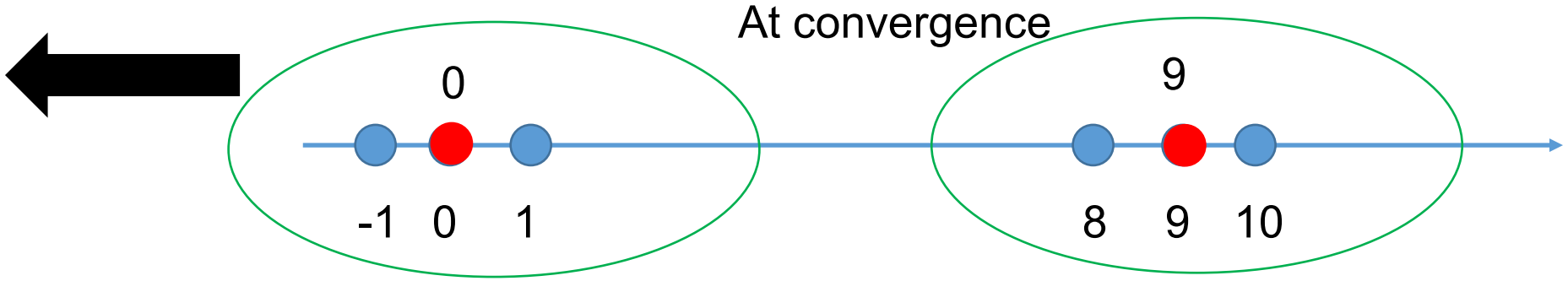
1. Step 1 
2. Step 2

Q2-3: Given the following data points in 1D: x1 = -1, x2 = 0, x3 = 1, x4 = 8, x5 = 9, x6 = 10, what is the distortion of x6 and the whole dataset respectively **at convergence?** Assume K=2 and we start with cluster centers c1 = 2 and c2 = 8.

1. 1, 0
2. 2, 2
3. 1, 4
4. 2, 4

Q2-3: Given the following data points in 1D: x1 = -1, x2 = 0, x3 = 1, x4 = 8, x5 = 9, x6 = 10, what is the distortion of x6 and the whole dataset respectively **at convergence?** Assume K=2 and we start with cluster centers c1 = 2 and c2 = 8.

1. 1, 0
2. 2, 2
3. 1, 4
4. 2, 4



Q2-4: If we choose number of clusters equal to number of data points, i.e. K = n, what will be the distortion of the dataset at convergence? Assume the starting cluster centers are same as the data points.

1. 0
2. n
3. 1
4. n-1

Q2-4: If we choose number of clusters equal to number of data points, i.e. K = n, what will be the distortion of the dataset at convergence? Assume the starting cluster centers are same as the data points.

1. 0 
2. n
3. 1
4. n-1

Q3-1: If we run K-means clustering twice with random starting cluster centers, are we guaranteed to get same clustering results?

1. Yes
2. No

Q3-1: If we run K-means clustering twice with random starting cluster centers, are we guaranteed to get same clustering results?

1. Yes
2. No 

Q3-2: Is it guaranteed that K-means will always terminate? Does K-means always lead to global optimum?

1. Yes, Yes
2. No, Yes
3. Yes, No
4. No, No

Q3-2: Is it guaranteed that K-means will always terminate? Does K-means always lead to global optimum?

1. Yes, Yes
2. No, Yes
3. Yes, No 
4. No, No

Q3-3: Which of the following could help for K-means to find a global optimum?

i) Run K-means only for a fixed number of iterations
ii) Run K-means multiple times with different starting cluster centers.
Iii) Pick the starting cluster centers intelligently.

1. only (i)
2. (i) and (ii)
3. (i) and (iii)
4. (ii) and (iii)

Q3-3: Which of the following could help for K-means to find a global optimum?

i) Run K-means only for a fixed number of iterations
ii) Run K-means multiple times with different starting cluster centers.
Iii) Pick the starting cluster centers intelligently.

1. only (i)
2. (i) and (ii)
3. (i) and (iii)
4. (ii) and (iii)