Q4

April 17, 2024

p1: (0,0),

p2: (1,2),

p3: [2,0),

p4: [8,8),

p5: [9,9),

p6: [10,8)

choose two initial centroid $\mu_1 : (1, 1)$ and $\mu_2 : (9, 9)$

0.1 Iteration 1

computer the distance of poinst to the centrol

	C_1	C_2	k
p1	$\sqrt{(0-1)^2+(0-1)^2} =$ 1.41	$\sqrt{(0-9)^2+(0-9)^2} =$ 12.7	C_1
p2	$\sqrt{(1-1)^2+(2-1)^2} =$ 1.00	$\sqrt{(1-9)^2+(2-9)^2} =$ 10.6	C_1
p3	$\sqrt{(2-1)^2+(0-1)^2} =$ 1.41	$\sqrt{(2-9)^2+(0-9)^2} =$ 11.4	C_1
p4	$\sqrt{(8-1)^2+(8-1)^2} =$ 9.90	$\sqrt{(8-9)^2+(8-9)^2} =$ 1.41	C_2
p5	$\sqrt{(9-1)^2+(9-1)^2} =$ 11.3	$\sqrt{(9-9)^2+(9-9)^2} =$ 0.00	C_2
p6	$\sqrt{(10-1)^2+(8-1)^2} =$ 11.4	$\sqrt{(10-9)^2+(8-9)^2} =$ 1.41	C_2

optimization of centroids

$$\mu_1 = \frac{\sum_n r_{n1} x_n}{\sum_n r_{n1}}$$

$$= \frac{1 \cdot [0,0] + 1 \cdot [1,2] + 1 \cdot [2,0] + 0 \cdot [8,8] + 0 \cdot [9,9] + 0 \cdot [10,8]}{1+1+1+0+0+0}$$

$$[1, \frac{2}{3}]$$

$$\mu_2 = \frac{\sum_n r_{n2} x_n}{\sum_n r_{n2}}$$
$$= \frac{0 \cdot [0,0] + 0 \cdot [1,2] + 0 \cdot [2,0] + 1 \cdot [8,8] + 1 \cdot [9,9] + 1 \cdot [10,8]}{0+0+0+1+1+1}$$
$$[9, \frac{25}{3}]$$

0.2 Iteration 2

computer the distance of poinst to the centrol

	C_1	C_2	k
p1	$\sqrt{(0-1)^2+(0-\frac{2}{3})^2} =$ 1.20	$\sqrt{(0-9)^2+(0-\frac{25}{3})^2} =$ 12.3	C_1
p2	$\sqrt{(1-1)^2+(2-\frac{2}{3})^2} =$ 1.33	$\sqrt{(1-9)^2+(2-\frac{25}{3})^2} =$ 10.2	C_1
p3	$\sqrt{(2-1)^2+(0-\frac{2}{3})^2} =$ 1.20	$\sqrt{(2-9)^2+(0-\frac{25}{3})^2} =$ 10.9	C_1
p4	$\sqrt{(8-1)^2+(8-\frac{2}{3})^2} =$ 10.1	$\sqrt{(8-9)^2+(8-\frac{25}{3})^2} =$ 1.05	C_2
p5	$\sqrt{(9-1)^2+(9-\frac{2}{3})^2} =$ 11.6	$\sqrt{(9-9)^2+(9-\frac{25}{3})^2} =$ 0.67	C_2
p6	$\sqrt{(10-1)^2+(8-\frac{2}{3})^2} =$ 11.6	$\sqrt{(10-9)^2+(8-\frac{25}{3})^2} =$ 1.05	C_2

optimization of centroids

$$\mu_1 = \frac{\sum_n r_{n1} x_n}{\sum_n r_{n1}}$$
$$= \frac{1 \cdot [0,0] + 1 \cdot [1,2] + 1 \cdot [2,0] + 0 \cdot [8,8] + 0 \cdot [9,9] + 0 \cdot [10,8]}{1+1+1+0+0+0}$$
$$[1, \frac{2}{3}]$$

$$\mu_2 = \frac{\sum_n r_{n2} x_n}{\sum_n r_{n2}}$$
$$= \frac{0 \cdot [0,0] + 0 \cdot [1,2] + 0 \cdot [2,0] + 1 \cdot [8,8] + 1 \cdot [9,9] + 1 \cdot [10,8]}{0+0+0+1+1+1}$$

$$[19, \frac{25}{3}]$$

The centroid has converge

[2]:
```
from sklearn.cluster import KMeans
```

[5]:
```
X = [(0,0), (1,2), (2,0), (8,8), (9,9), (10,8)]
```

[6]:
```
c = 2  # number of clusters
kmeans = KMeans(n_clusters=c, random_state=42)
kmeans = kmeans.fit(X)

print(kmeans.cluster_centers_)
```

```
[[1.         0.66666667]
 [9.         8.33333333]]
```