Given data points:

$x^{(1)} = (2, 8)$, $x^{(2)} = (2, 5)$, $x^{(3)} = (1, 2)$, $x^{(4)} = (5, 8)$

$x^{(5)} = (7, 3)$, $x^{(6)} = (6, 4)$, $x^{(7)} = (8, 4)$, $x^{(8)} = (4, 7)$

Distance Matrix:

| | $x^{(1)}$ | $x^{(2)}$ | $x^{(3)}$ | $x^{(4)}$ | $x^{(5)}$ | $x^{(6)}$ | $x^{(7)}$ | $x^{(8)}$ |
|---|---|---|---|---|---|---|---|---|
| $x^{(1)}$ | 0 | 3.0000 | 6.0828 | 3.0000 | 7.0711 | 5.6569 | 7.2111 | 2.2361 |
| $x^{(2)}$ | 3.0000 | 0 | 3.1623 | 4.2426 | 5.3852 | 4.1231 | 6.0828 | 2.8284 |
| $x^{(3)}$ | 6.0828 | 3.1623 | 0 | 7.2111 | 6.0828 | 5.3852 | 7.2801 | 5.8310 |
| $x^{(4)}$ | 3.0000 | 4.2426 | 7.2111 | 0 | 5.3852 | 4.1231 | 5.0000 | 1.4142 |
| $x^{(5)}$ | 7.0711 | 5.3852 | 6.0828 | 5.3852 | 0 | 1.4142 | 1.4142 | 5.0000 |
| $x^{(6)}$ | 5.6569 | 4.1231 | 5.3852 | 4.1231 | 1.4142 | 0 | 2.0000 | 3.6056 |
| $x^{(7)}$ | 7.2111 | 6.0828 | 7.2801 | 5.0000 | 1.4142 | 2.0000 | 0 | 5.0000 |
| $x^{(8)}$ | 2.2361 | 2.8284 | 5.8310 | 1.4142 | 5.0000 | 3.6056 | 5.0000 | 0 |

Based on the distance Matrix the cloest centroids:

$$c^{(1)} = c^{(4)} = c^{(8)} = 2$$
$$c^{(2)} = c^{(3)} = 1$$
$$c^{(5)} = c^{(6)} = c^{(7)} = 3$$

Initialization:

$$\mu^{(1)} = x^{(3)}$$
$$\mu^{(2)} = x^{(4)}$$
$$\mu^{(3)} = x^{(6)}$$

[Given $x^{(3)}, x^{(4)}$ & $x^{(6)}$ are initialisation cluster centers]

Move centroids:

$$\mu^{(1)} = \frac{1}{2}\left(x^{(2)} + x^{(3)}\right)$$

$$= \frac{1}{2}\left((2,5) + (1,2)\right)$$

$$= \cancel{\begin{pmatrix} 1.5 \\ 3.5 \end{pmatrix}} \; (1.5,\ 3.5)$$

$$\mu^{(2)} = \frac{1}{3}\left(x^{(1)} + x^{(4)} + x^{(8)}\right)$$

$$= \frac{1}{3}\left((2,8) + (5,8) + (4,7)\right)$$

$$= (3.67,\ 7.67)$$

$$\mu^{(3)} = \frac{1}{3}\left(x^{(5)} + x^{(6)} + x^{(7)}\right)$$

$$= \frac{1}{3}\left((7,3) + (6,4) + (8,4)\right)$$

$$= (7,\ 3.67)$$

Loss function: $J(c^{(1)}, \ldots, c^{(m)}, \mu_1, \ldots, \mu_K) = \frac{1}{m}\sum_{i=1}^{m} \|x^{(i)} - \mu_{c(i)}\|^2$

Before using K-means:

For $x^{(2)}$ and $x^{(3)}$, $c = x^{(3)}$

So, $\|x^{(2)} - x^{(3)}\|^2 + \|x^{(3)} - x^{(3)}\|^2$

$$= 3.1623^2 + 0 = 3.1623^2$$

For $x^{(1)}$, $x^{(4)}$, and $x^{(8)}$, $c = x^{(4)}$

So, $\|x^{(1)} - x^{(4)}\|^2 + \|x^{(4)} - x^{(4)}\|^2 + \|x^{(8)} - x^{(4)}\|^2$

$$= 3^2 + 0 + 1.4142^2 = 3^2 + 1.4142^2$$

For $x^{(5)}$, $x^{(6)}$ and $x^{(7)}$, $c = x^{(6)}$

So, $\|x^{(5)} - x^{(6)}\|^2 + \|x^{(6)} - x^{(6)}\|^2 + \|x^{(7)} - x^{(6)}\|^2$

$$= 1.4142^2 + 0^2 + 2^2$$

$$= 1.4142^2 + 2^2$$

$$\therefore J_0 = \frac{1}{8}\left(3.1623^2 + 3^2 + 1.4142^2 + 1.4142^2 + 2^2\right)$$

$$= 3.375.$$

After 1'st Iteration of k-means:

$$\mu^{(1)} = (1.5, 3.5) \; ; \; \mu^{(2)} = (3.67, 7.67)$$

$$\mu^{(3)} = (7, 3.67)$$

For $x^{(2)}$ and $x^{(3)}$, $\mu_c = (1.5, 3.5)$

So, $\|x^{(2)} - \mu_c\|^2 + \|x^{(3)} - \mu_c\|^2$

$$= 2.5 + 2.5 = 5$$

For $x^{(1)}$, $x^{(4)}$ and $x^{(8)}$, $\mu_c = (3.67, 7.67)$

So, $\|x^{(1)} - \mu_c\|^2 + \|x^{(4)} - \mu_c\|^2 + \|x^{(8)} - \mu_c\|^2$

$$= 2.8978 + 1.8778 + 0.5578$$

$$= 5.3334$$

For $x^{(5)}$, $x^{(6)}$ and $x^{(7)}$, $\mu_c = (7, 3.67)$

So, $\|x^{(5)} - \mu_c\|^2 + \|x^{(6)} - \mu_c\|^2 + \|x^{(7)} - \mu_c\|^2$

$$= 0.4489 + 1.1089 + 1.1089$$

$$= 2.6667.$$

$$\therefore J_1 = \frac{1}{8}\left(5 + 5.3334 + 2.6667\right)$$

$$= 1.6250$$

## 2nd Iteration:

Now, $\mu^{(1)} = (1.5, 3.5)$

$\mu^{(2)} = (3.67, 7.67)$

$\mu^{(3)} = (7, 3.67)$

Distance:

| | $\mu^{(1)}$ | $\mu^{(2)}$ | $\mu^{(3)}$ |
|---|---|---|---|
| $x^{(1)}$ | 4.52 | 1.70 | 6.61 |
| $x^{(2)}$ | 1.58 | 3.15 | 5.17 |
| $x^{(3)}$ | 1.58 | 6.27 | 6.23 |
| $x^{(4)}$ | 5.70 | 1.37 | 4.77 |
| $x^{(5)}$ | 5.52 | 5.79 | 0.67 |
| $x^{(6)}$ | 4.52 | 4.35 | 1.05 |
| $x^{(7)}$ | 6.52 | 5.68 | 1.05 |
| $x^{(8)}$ | 4.30 | 0.74 | 4.48 |

Closest centroids:

$c^{(2)} = c^{(3)} = 1$

$c^{(1)} = c^{(4)} = c^{(8)} = 2$

$c^{(5)} = c^{(6)} = c^{(7)} = 3$

Move centroids:

$$\mu^{(1)} = \frac{1}{2}\left(x^{(2)} + x^{(3)}\right)$$

$$= (1.5, 3.5)$$

$$\mu^{(2)} = \frac{1}{3}\left(x^{(1)} + x^{(4)} + x^{(8)}\right)$$

$$= (3.67, 7.67)$$

$$\mu^{(3)} = \frac{1}{3}\left(x^{(5)} + x^{(6)} + x^{(7)}\right)$$

$$= (7, 3.67)$$

Loss Function in 2nd Iteration: $\left[ \frac{1}{m} \sum_{i=1}^{m} \| x^{(i)} - \mu_{c^{(i)}} \|^2 \right]$

For $x^{(2)}$ and $x^{(3)}$, $\mu_c = (1.5, \cancel{5}\,3.5)$

So, $\| x^{(2)} - \mu_c \|^2 + \| x^{(3)} - \mu_c \|^2$

$= 2.5 + 2.5 = 5$

For $x^{(1)}$, $x^{(4)}$, and $x^{(8)}$, $\mu_c = (3.67, 7.67)$

So, $\| x^{(1)} - \mu_c \|^2 + \| x^{(4)} - \mu_c \|^2 + \| x^{(8)} - \mu_c \|^2$

$= 2.8978 + 1.8778 + 0.5578$

$= 5.3334$

For $x^{(5)}$, $x^{(6)}$ and $x^{(7)}$, $\mu_c = (7, 3.67)$

So, $\| x^{(5)} - \mu_c \|^2 + \| x^{(6)} - \mu_c \|^2 + \| x^{(7)} - \mu_c \|^2$

$= 0.4489 + 1.1089 + 1.1089$

$= 2.6667.$

$\therefore J_2 = \frac{1}{8}(5 + 5.3334 + 2.6667)$

$= 1.6250.$