K-Means Clustering Example

We have the following data points:

$$x^{(1)} = (1, 1)$$
$$x^{(2)} = (1, 2)$$
$$x^{(3)} = (2, 2)$$
$$x^{(4)} = (5, 5)$$
$$x^{(5)} = (6, 5)$$
$$x^{(6)} = (6, 6)$$

We want to group them into 2 clusters ($K = 2$).

Step 1: Initialization

Randomly initialize 2 cluster centroids:

$$\mu_1 = (1, 1)$$
$$\mu_2 = (5, 5)$$

Step 2: Assignment Step

Assign each point to the nearest centroid.

Point $x^{(1)} = (1, 1)$: Closest to μ_1.

Point $x^{(2)} = (1, 2)$:

$$\text{Distance to } \mu_1 : \|x^{(2)} - \mu_1\|^2 = (1-1)^2 + (2-1)^2 = 1$$
$$\text{Distance to } \mu_2 : \|x^{(2)} - \mu_2\|^2 = (1-5)^2 + (2-5)^2 = 25$$

Closest to μ_1.

Point $x^{(3)} = (2, 2)$:

$$\text{Distance to } \mu_1 : \|x^{(3)} - \mu_1\|^2 = (2-1)^2 + (2-1)^2 = 2$$
$$\text{Distance to } \mu_2 : \|x^{(3)} - \mu_2\|^2 = (2-5)^2 + (2-5)^2 = 18$$

Closest to μ_1.

Point $x^{(4)} = (5, 5)$: Closest to μ_2.

Point $x^{(5)} = (6, 5)$**:**

$$\begin{aligned}&\text{Distance to } \mu_1 : \|x^{(5)} - \mu_1\|^2 = (6-1)^2 + (5-1)^2 = 41\\&\text{Distance to } \mu_2 : \|x^{(5)} - \mu_2\|^2 = (6-5)^2 + (5-5)^2 = 1\\&\ \text{Closest to } \mu_2.\end{aligned}$$

Point $x^{(6)} = (6, 6)$**:**

$$\begin{aligned}&\text{Distance to } \mu_1 : \|x^{(6)} - \mu_1\|^2 = (6-1)^2 + (6-1)^2 = 50\\&\text{Distance to } \mu_2 : \|x^{(6)} - \mu_2\|^2 = (6-5)^2 + (6-5)^2 = 2\\&\ \text{Closest to } \mu_2.\end{aligned}$$

Cluster Assignments:

$$c(1) = 1, \quad c(2) = 1, \quad c(3) = 1, \quad c(4) = 2, \quad c(5) = 2, \quad c(6) = 2.$$

Step 3: Update Step

Calculate the new centroids by taking the average of the points in each cluster.

New μ_1: Average of $x^{(1)}, x^{(2)}, x^{(3)}$

$$\mu_1 = \left(\frac{1+1+2}{3}, \frac{1+2+2}{3}\right) = \left(\frac{4}{3}, \frac{5}{3}\right) \approx (1.33, 1.67).$$

New μ_2: Average of $x^{(4)}, x^{(5)}, x^{(6)}$

$$\mu_2 = \left(\frac{5+6+6}{3}, \frac{5+5+6}{3}\right) = \left(\frac{17}{3}, \frac{16}{3}\right) \approx (5.67, 5.33).$$

Step 4: Repeat

Repeat the assignment and update steps with the new centroids.

Cluster Assignments:

$$c(1) = 1, \quad c(2) = 1, \quad c(3) = 1, \quad c(4) = 2, \quad c(5) = 2, \quad c(6) = 2.$$

Since the cluster assignments did not change, the algorithm converges, and we have our final clusters.

Final Clusters

- Cluster 1: $x^{(1)}, x^{(2)}, x^{(3)}$
- Cluster 2: $x^{(4)}, x^{(5)}, x^{(6)}$

Optimization Objective

The optimization objective J is the average squared distance between each point and its assigned centroid. For this example:

$$J(c(1), \ldots, c(6), \mu_1, \mu_2) = \frac{1}{6} \sum_{i=1}^{6} \|x^{(i)} - \mu_{c(i)}\|^2.$$

Calculating each term:

$$\begin{aligned}
\|x^{(1)} - \mu_1\|^2 &= (1 - 1.33)^2 + (1 - 1.67)^2 = 0.1089 + 0.4489 = 0.5578 \\
\|x^{(2)} - \mu_1\|^2 &= (1 - 1.33)^2 + (2 - 1.67)^2 = 0.1089 + 0.1089 = 0.2178 \\
\|x^{(3)} - \mu_1\|^2 &= (2 - 1.33)^2 + (2 - 1.67)^2 = 0.4489 + 0.1089 = 0.5578 \\
\|x^{(4)} - \mu_2\|^2 &= (5 - 5.67)^2 + (5 - 5.33)^2 = 0.4489 + 0.1089 = 0.5578 \\
\|x^{(5)} - \mu_2\|^2 &= (6 - 5.67)^2 + (5 - 5.33)^2 = 0.1089 + 0.1089 = 0.2178 \\
\|x^{(6)} - \mu_2\|^2 &= (6 - 5.67)^2 + (6 - 5.33)^2 = 0.1089 + 0.4489 = 0.5578.
\end{aligned}$$

$$J = \frac{1}{6}(0.5578 + 0.2178 + 0.5578 + 0.5578 + 0.2178 + 0.5578) = \frac{1}{6}(2.6668) \approx 0.4445$$