Fuzzy C-Means Clustering

March 2025

1 Introduction

Fuzzy C-Means (FCM) clustering is a method of clustering that allows one piece of data to belong to more than one cluster. This is in contrast to traditional clustering methods like K-Means, where each data point is assigned to only one cluster. FCM is based on the concept of fuzzy sets, where each data point has a degree of belonging to each cluster rather than belonging exclusively to one cluster.

1.1 Definition of Fuzzy C-Means

The Fuzzy C-Means algorithm divides a set of data points into a specified number C of clusters. Each data point is assigned a membership value for each cluster, indicating the degree of association between the point and each cluster. The goal of FCM is to minimize the following objective function:

$$J(U, V) = \sum_{i=1}^{n} \sum_{c=1}^{C} (\mu_{ic})^m \cdot \|x_i - v_c\|^2$$

where:

- n is the number of data points,
- C is the number of clusters,
- μ_{ic} is the membership degree of data point x_i in cluster c,
- v_c is the center of cluster c,
- $\| \cdot \|$ is the Euclidean distance, and
- m is the fuzziness parameter, controlling the degree of fuzziness (typically $m > 1$).

The algorithm minimizes this objective function by adjusting the membership values μ_{ic} and the cluster centers v_c iteratively.

1.2 How Fuzzy C-Means Differs from K-Means

The key difference between Fuzzy C-Means and K-Means lies in how data points are assigned to clusters:

- **K-Means Clustering:** In K-Means, each data point is assigned to exactly one cluster, based on its proximity to the nearest cluster center. The assignment is crisp, meaning a point either belongs to a cluster or it does not.
- **Fuzzy C-Means Clustering:** In contrast, FCM assigns each data point a degree of membership to every cluster. A data point can belong to multiple clusters with varying degrees of membership. This is particularly useful when the data has inherent overlapping characteristics.

The fuzziness parameter m in FCM controls how soft or hard the assignments are. When m is set to 1, the algorithm reduces to a crisp classification like K-Means. As m increases, the algorithm becomes fuzzier, allowing for more overlap between clusters.

2 Fuzzy C-Means Clustering Algorithm: Step-by-Step Explanation

2.1 Step 1: Define the Problem

We have four data points in 2D space:

$$x_1 = (1, 3), \quad x_2 = (1.5, 3.2), \quad x_3 = (1.3, 2.8), \quad x_4 = (3, 1)$$

Here, $x_k = (x_{k1}, x_{k2})$ represents the k-th data point in a 2-dimensional space, where x_{k1} and x_{k2} are the first and second features of the point.

2.2 Step 2: Set Algorithm Parameters

- Number of Clusters (C): $C = 2$. This is the number of clusters to form.
- Fuzziness Parameter (m): $m = 2$. This controls the degree of overlap between clusters.
- Convergence Criterion (ϵ): $\epsilon = 0.01$. This is the threshold for stopping the algorithm.

2.3 Step 3: Initialize the Membership Matrix

We initialize the membership matrix U_0 as:

$$U_0 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In this initial matrix:

- The first three points belong to Cluster 1.
- The fourth point belongs to Cluster 2.

2.4 Step 4: Calculate Cluster Centers

The cluster centers V are calculated using the formula:

$$V_{ij} = \frac{\sum_{k=1}^{n} (\mu_{ik})^m \cdot x_{kj}}{\sum_{k=1}^{n} (\mu_{ik})^m}$$

where:

- μ_{ik} is the membership of the k-th data point in the i-th cluster raised to the power m,
- x_{kj} is the j-th feature of the k-th data point,
- n is the total number of data points.

2.4.1 Cluster 1 Center (V_1)

$$V_{11} = \frac{(1^2 \cdot 1) + (1^2 \cdot 1.5) + (1^2 \cdot 1.3) + (0^2 \cdot 3)}{1^2 + 1^2 + 1^2 + 0^2} = \frac{1 + 1.5 + 1.3 + 0}{3} = 1.26$$

$$V_{12} = \frac{(1^2 \cdot 3) + (1^2 \cdot 3.2) + (1^2 \cdot 2.8) + (0^2 \cdot 1)}{1^2 + 1^2 + 1^2 + 0^2} = \frac{3 + 3.2 + 2.8 + 0}{3} = 3.0$$

So, $V_1 = (1.26, 3.0)$.

2.4.2 Cluster 2 Center (V_2)

$$V_{21} = \frac{(0^2 \cdot 1) + (0^2 \cdot 1.5) + (0^2 \cdot 1.3) + (1^2 \cdot 3)}{0^2 + 0^2 + 0^2 + 1^2} = \frac{0 + 0 + 0 + 3}{1} = 3$$

$$V_{22} = \frac{(0^2 \cdot 3) + (0^2 \cdot 3.2) + (0^2 \cdot 2.8) + (1^2 \cdot 1)}{0^2 + 0^2 + 0^2 + 1^2} = \frac{0 + 0 + 0 + 1}{1} = 1$$

So, $V_2 = (3, 1)$.

2.5 Step 5: Calculate Distances

Compute the Euclidean distance of each data point from the cluster centers.

2.5.1 Distance Formula

$$d_{ik} = \sqrt{(x_{kj} - V_{ij})^2}$$

where x_{kj} is the j-th feature of the k-th data point, and V_{ij} is the j-th feature of the center of cluster i.

2.5.2 Calculations

- Distance from $x_1 = (1, 3)$:

$$d_{11} = \sqrt{(1 - 1.26)^2 + (3 - 3)^2} = \sqrt{(-0.26)^2 + 0} = 0.26$$

$$d_{21} = \sqrt{(1 - 3)^2 + (3 - 1)^2} = \sqrt{(-2)^2 + (2)^2} = \sqrt{4 + 4} = 2.82$$

- Distance from $x_2 = (1.5, 3.2)$:

$$d_{12} = \sqrt{(1.5 - 1.26)^2 + (3.2 - 3)^2} = \sqrt{(0.24)^2 + (0.2)^2} = \sqrt{0.0576 + 0.04} = 0.31$$

$$d_{22} = \sqrt{(1.5 - 3)^2 + (3.2 - 1)^2} = \sqrt{(-1.5)^2 + (2.2)^2} = \sqrt{2.25 + 4.84} = 2.66$$

- Distance from $x_3 = (1.3, 2.8)$:

$$d_{13} = \sqrt{(1.3 - 1.26)^2 + (2.8 - 3)^2} = \sqrt{(0.04)^2 + (-0.2)^2} = \sqrt{0.0016 + 0.04} = 0.20$$

$$d_{23} = \sqrt{(1.3 - 3)^2 + (2.8 - 1)^2} = \sqrt{(-1.7)^2 + (1.8)^2} = \sqrt{2.89 + 3.24} = 2.47$$

- Distance from $x_4 = (3, 1)$:

$$d_{14} = \sqrt{(3 - 1.26)^2 + (1 - 3)^2} = \sqrt{(1.74)^2 + (-2)^2} = \sqrt{3.0276 + 4} = 2.65$$

$$d_{24} = \sqrt{(3 - 3)^2 + (1 - 1)^2} = \sqrt{0 + 0} = 0$$

2.6 Step 6: Update Membership Matrix

The updated membership values are calculated using:

$$\mu_{ik}^{r+1} = \left(\sum_{j=1}^{C} \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}} \right)^{-1}$$

2.6.1 Calculations

- For $x_1 = (1, 3)$:

$$\mu_{11} = \left(\left(\frac{0.26}{0.26} \right)^2 + \left(\frac{0.26}{2.82} \right)^2 \right)^{-1} = (1 + 0.0085)^{-1} = 0.991$$

$$\mu_{21} = 1 - \mu_{11} = 1 - 0.991 = 0.009$$

- For $x_2 = (1.5, 3.2)$:

$$\mu_{12} = \left(\left(\frac{0.31}{0.31} \right)^2 + \left(\frac{0.31}{2.66} \right)^2 \right)^{-1} = (1 + 0.0138)^{-1} = 0.982$$

$$\mu_{22} = 1 - \mu_{12} = 1 - 0.982 = 0.018$$

- For $x_3 = (1.3, 2.8)$:

$$\mu_{13} = \left(\left(\frac{0.20}{0.20}\right)^2 + \left(\frac{0.20}{2.47}\right)^2\right)^{-1} = (1 + 0.0325)^{-1} = 0.993$$

$$\mu_{23} = 1 - \mu_{13} = 1 - 0.993 = 0.007$$

- For $x_4 = (3, 1)$:

$$\mu_{14} = \left(\left(\frac{2.65}{2.65}\right)^2 + \left(\frac{2.65}{0}\right)^2\right)^{-1} = (1 + \infty)^{-1} = 0$$

$$\mu_{24} = 1 - \mu_{14} = 1 - 0 = 1$$

2.7 Step 7: Repeat Steps 4-6

We repeat Steps 4-6 until the membership matrix converges, i.e., when the change in membership values is smaller than the convergence criterion ϵ.