K-means clustering

distance : $\boxed{\text{dist}(a, b) = \| a - b \|_2^2}$

assigning to cluster j : chosen $j = \min\limits_{j \in [1, K]} \| x_i - z_j \|_2^2$

cost : $\boxed{\sum_{j=1}^{K} \sum_{i \in C_j} \| x_i - z_j \|_2^2}$

goal : $\min\limits_{z} \sum_{j=1}^{K} \sum_{i \in C_j} \| x_i - z_j \|_2^2$

new representatives:

$$\frac{d}{dz_j} \sum_{i \in C_j} \| x_i - z_j \|_2^2 = \sum_{i \in C_j} 2(x_i - z_j)$$

$$\frac{d}{dz_j} \text{cost} = 0 \implies \sum_{i \in C_j} (x_i - z_j) = 0$$

$$\sum_{i \in C_j} x_i - \underbrace{\sum_{i \in C_j} z_j}_{|C_j| \cdot z_j} = 0$$

$$|C_j| \cdot z_j = \sum_{i \in C_j} x_i$$

$$\boxed{z_j = \frac{\sum_{i \in C_j} x_i}{|C_j|}}$$

example

A (2, 3)
B (3, 3)
C (6, 5)
D (8, 8)

1. specify K → $K = 2$

2. randomly select representatives $\{z_1 \dots z_k\}$

A (2, 3) → z_1 (2, 3)
B (3, 3)
C (6, 5) → z_2 (6, 5)
D (8, 8)

3. assign data points to clusters

$\min_{j \in [1,k]} \text{dist}(x_i, z_j)$

point	$\text{dist}(x_i, z)$	$\text{dist}(x_i, z_2)$	assigned cluster
A (2, 3)	0	20	C_1
B (3, 3)	1	13	C_1
C (6, 5)	20	0	C_2
D (8, 8)	61	13	C_2

4 update representatives

$$z_j = \frac{\sum_{i \in C_j} x_i}{|C_j|} \rightarrow \text{center of cluster}$$

$$z_1 = \frac{(2,3) + (3,3)}{2} = (2.5, 3)$$

$$z_2 = \frac{(6,5) + (8,8)}{2} = (7, 6.5)$$

5 iterate until minimal changes

Kernel K-NNs

mapping x to a higher dimension $x = \phi(x)$

$$\Rightarrow \quad z_j = \frac{\sum_{i \in C_j} \phi(x_i)}{|C_j|}$$

calculating distance:

$$\| \phi(x_i) - z_j \|_2^2 = \phi(x_i)^T \phi(x_i) - 2\, \phi(x_i)^T \cdot z_j + z_j^T z_j$$

$$= \phi(x_i)^T \phi(x_i) - 2\, \phi(x_i)^T \frac{\sum_{t \in C_j} \phi(x_t)}{|C_j|} + \left(\frac{\sum_{t \in C_j} \phi(x_t)}{|C_j|}\right)^T \cdot \frac{\sum_{t \in C_j} \phi(x_t)}{|C_j|}$$

$$= \phi(x_i)^T \phi(x_i) - 2 \frac{\sum_{t \in C_j} \phi(x_i)^T \phi(x_t)}{|C_j|} + \frac{\sum_{t \in C_j} \phi(x_t)^T \phi(x_t)}{|C_j|^2}$$

Kernel trick!

polynomial $\rightarrow \phi(a)^T \phi(b) = (a^T b^T + 1)^d$

RBF $\rightarrow \phi(a)^T \phi(b) = \exp\left(\gamma \| a - b \|_2^2\right)$