k-means Algorithm: Decoding Step

encoder/assignment is known, what is the best decoder/codebook?

Optimal solution for the decoding step

$$x_i \longrightarrow [encoder] \longrightarrow index \ y(x_i) \longrightarrow [decoder] \longrightarrow centroid \ z_{y(x_i)}$$

second step: decoder which minimise the training reconstruction error

$$J(\mathcal{C}) = \frac{1}{n} \sum_{i=1}^{n} d(\mathbf{x}_i, \mathbf{z}_{y(\mathbf{x}_i)})^2 = \sum_{j=1}^{k} \left(\frac{|\mathcal{C}_j|}{n} \right) \left(\frac{1}{|\mathcal{C}_j|} \sum_{i \in \mathcal{C}_j} d(\mathbf{x}_i, \mathbf{z}_{y(\mathbf{x}_i)})^2 \right)$$

solution: move centroid z_i to the center of gravity of cluster C_i

$$\mathbf{z}_j = \frac{1}{|\mathcal{C}_j|} \sum_{i \in \mathcal{C}_i} \mathbf{x}_i$$