k-means Algorithm: Encoding Step

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

Optimal solution for the encoding step

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

first step: encoder which minimise the training reconstruction error

$$J(C) = \frac{1}{n} \sum_{i=1}^{n} d(\mathbf{x}_i, \mathbf{z}_{y(\mathbf{x}_i)})^2$$

solution: assign x_i to the closest centroid $z_{y(x_i)}$

$$y(\mathbf{x}_i) = \arg\min_{j=1...k} d(\mathbf{x}_i, \mathbf{z}_j)$$