squared Euclidean Distance

$$W(C_k) = \frac{1}{|C_k|} \sum_{i,i' \in C_k} \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2$$

C_k: number of observation in the kth observation

$$\frac{1}{|C_k|} \sum_{i,i' \in C_k} \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2 = 2 \sum_{i \in C_k} \sum_{j=1}^{p} (x_{ij} - \bar{x}_{kj})^2$$

$$\bar{x}_{kj} = \frac{1}{|C_k|} \sum_{i \in C_k} x_{ij}$$

$$\Rightarrow \frac{1}{|C_k|} \sum_{i,i' \in C_k} \sum_{j=1}^{p} ((x_{ij} - \bar{x}_{kj}) - (x_{i'j} - \bar{x}_{kj}))^2$$

$$\Rightarrow \frac{1}{|C_k|} \sum_{i,i' \in C_k} \sum_{j=1}^{p} ((x_{ij} - \bar{x}_{kj})^2 - 2(x_{ij} - \bar{x}_{kj})(x_{i'j} - \bar{x}_{kj}) + (x_{i'j} - \bar{x}_{kj})^2)$$

$$\Rightarrow \frac{|C_k|}{|C_k|} \sum_{i \in C_k} \sum_{j=1}^{p} (x_{ij} - \bar{x}_{kj})^2 + \frac{|C_k|}{|C_k|} \sum_{i' \in C_k} \sum_{j=1}^{p} (x_{i'j} - \bar{x}_{kj})^2 - \frac{2}{|C_k|} \sum_{i,i' \in C_k} \sum_{j=1}^{p} (x_{ij} - \bar{x}_{kj})(x_{i'j} - \bar{x}_{kj})$$

$$= 2 \sum_{i \in C_k} \sum_{j=1}^{p} (x_{ij} - \bar{x}_{kj})^2 + 0$$

#

2.

$$A \quad \frac{1}{|C_k|} \sum_{i,i' \in C_k} \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2 \Rightarrow \text{Squared Euclidean Dis}$$

Within cluster variation

$$B = \sum_{i \in C_k} \sum_{j=1}^{p} (x_{ij} - \bar{x}_{kj})^2$$

$\bar{x}_{kj}$: mean of feature j in cluster C_k

In algorithm Step 2(a), B can be minimized since the centroid is the new mean for each cluster.

Therefore if B can be minimized, then, A (squared Euclidean Dis. in each cluster) can be minimized, too.