

index.C (clusterSim)

Hubert & Levin C index (internal cluster quality index)

$$C(u) = \frac{D(u) - D_{min}}{D_{max} - D_{min}}$$

where: $D(u)$ – all within-cluster dissimilarities in a partition of the objects into u clusters (the partition has a total of r such dissimilarities),

D_{min} – the sum of the r smallest dissimilarities in distance matrix,

D_{max} – the sum of the r largest dissimilarities in distance matrix,

u – number of clusters.

The value of u , which minimizes $C(u)$, is regarded as specifying the number of clusters.

References

Hubert, L.J., Levin, J.R. (1976), A General Statistical Framework for Assessing Categorical Clustering in Free Recall, *Psychological Bulletin*, Vol. 83, No. 6, 1072-1080.