How many Clusters: A Validation Index for arbitrary shaped clusters.
A. Baya & P.M. Granitto

Additional file 6: Multidimensional Scaling

This additional file includes the output of multidimensional scaling (MDS) for three artificial datasets, Three Ovals (Figure 1.a), Two Bars (Figure 2), and Three Rings (Figure 3.b). Two projections are shown in each case, one obtained with the Euclidean distance and other with the Maximum Edge Distance (Equation 2). The first two cases corresponds to the Three Ovals and Two Bars dataset pictured in Figure 12.a and Figure 13.a. There are two versions of the Two Bars Dataset, Figure 13.a where both clusters are separated at least by the maximum 1-nn distance and Figure 13.d where both clusters are separated by less than the maximum 1-nn distance. Figure 12.b and Figure 12.c show the first two components from multidimensional scaling obtained with Euclidean and Maximum Edge Distance respectively for the Three Ovals Dataset. Similarly Figure 13.b and Figure 13.c correspond to the projections of the Euclidean Distance and Maximum Edge Distance for the Two Bars Dataset while Figure 13.e and Figure 13.f correspond again to the projections of Two Bars but with small separation (Figure 12.c). Finally, Figure 14.a shows the Three Rings dataset though in this case the first three components form multidimensional scaling are shown. Figure 14.b,c are the projections corresponding to Euclidean Distance while Figure 14.d,e correspond to the Maximum Edge Distance.

Multidimensional scaling applied the Euclidean metric has a known interpretation it was included to contrast both results since Euclidean is the base metric from which MED is derived. The projections from Maximum Edge Distance shown in Figure 12.c, 13.c and Figure 14.d-e, form compact structures respect to the separation between clusters, although these structures may have arbitrary shapes in the Euclidean space. Section II-B (see Figure 2) shows that there is a limit to the detection of clusters, which can be clearly seen in Figure 13.f. The clusters from the previous figure do not appear compact in the projection since they are too close in the sense of first neighbors.

![Fig. 12. Three Ovals Dataset. a) Original Dataset. b) MDS of Euclidean’s distance first two components, x-axis corresponds to the first MDS component and y-axis to the second component. c) MDS of the Maximum Edge Distance first two components, x-axis corresponds to the first MDS component and y-axis to its second component.](image-url)
Fig. 13. Two versions of the Two Bars Dataset. One with good separation between clusters and the other with small separation between them. This set of figures shows the importance of this parameter under the MED distance. a) Original Dataset well separated. b) MDS first two components using Euclidean distance, x-axis corresponds to the first component and y-axis to the second component. Corresponds to the Dataset from point a). c) MDS first two components using MED distance, x-axis corresponds to the first MDS component and y-axis to the second component. Corresponds to the Dataset from point a). d) Original Dataset with small separation. e) MDS first two components using Euclidean distance, x-axis corresponds to the first MDS component and y-axis to the second component. Corresponds to the Dataset from point d). f) MDS first two components using MED distance, x-axis corresponds to the first MDS component and y-axis to the second component. Corresponds to the Dataset from point d).
Fig. 14. Three Rings Dataset. The first three components are needed in this case to show the separation between clusters using MED distance. 

(a) Three Rings Dataset

(b) MDS Euclidean Metric components one and two

(c) MDS Euclidean Metric components one and three

(d) MDS of MED Metric components one and two

(e) MDS of MED Metric components one and three

The first three components are needed in this case to show the separation between clusters using MED distance. a) MDS first two components using Euclidean distance, x-axis corresponds to the first MDS component and y-axis to the second component. b) MDS first and third components of Euclidean distance, x-axis corresponds to the first MDS component and y-axis to the third component. c) MDS first two components using MED distance, x-axis corresponds to the first MDS component and y-axis to the second component. d) MDS first and third components using MED distance, x-axis corresponds to the first MDS component and y-axis to the third component.