Dissimilarity Analysis for Classification

Adam Cardinal-Stakenas

28 February 2014

# The Dissimilarity Representation

- By computing dissimilarities between all pairs of observations, we get an easily interpretable representation of the data, even when the data comes from a hard-to-work-in space.

- The problem is that for a given type of data there may be many dissimilarities which might be applied.
- How do we choose one?

Choosing a Dissimilarity for Classification

- If we're trying to classify data, the within- and between- class distances of the training data are a useful indicator of class separation.
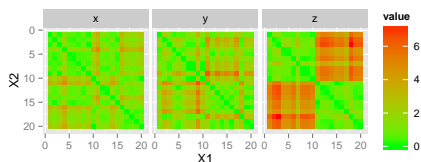
- We can use this to develop a framework for choosing a dissimilarity.