Trosset & Priebe, 2008

Abstract

out-of-sample stuff can happen on multilinear scaling, too

Introduction

Take some pairwise dissimilarity matrix for n objects $\Delta \in \mathbb{R}^{n\times n}$. If we want to classify these n objects, we can embed them into Euclidean space and then perform some classification procedure. A lot of the time, we need labels for this.

Suppose now that we have some oos objects, and we want to stick them into Euclidean space without re-embedding.

There are a few different criteria by which we could do this, like the method of standards and landmark MDS.

Classical multidimensional scaling

if the dissimilarity $[\delta_{ij}^2]$ in Δ is $||x_i - x_j||^2$, then Δ is a Euclidean distance matrix.

out-of-sample extension

if we add a vertex, we lose some properties of Δ, but we can get those properties back

Simulated data

Consider the Δ for four points arranged in a unit square. Now add a new point not on the square. the oos embedding preserves the square, but in-sample embedding methods don't.

Hippocampal dissimilarity data

we can do it with brains too

Discussion

don't think about this as a spectral technique, think about it as a least-squares technique with a spectral solution.