Robust embeddings – intrinsic dimensionality experiment

Abstract

Experiments indicate that the robust metric for embedding is much more informative for manifold learning.

We ran two ISOMAP-style experiments. For the first, we used 1000 images of handwritten ones from the MNIST database. We created the 7-nearest-neighbor graph on the dataset, then recorded the path distances between points. These path distances are conjectured to be reasonable approximations to geodesic distances. We compare the embeddings given by classical MDS and the SDP-based approach (SDS).

The ones dataset is conjectured to be a manifold of dimension 7 or 8 (Hein & Audibert, 2005). One would therefore hope that a plot of embedding error as a function of dimensionality would contain an elbow around those values. What we find is that the MDS cost does not contain such an elbow: rather, it slowly degrades. However, the robust cost contains a very obvious elbow.

The MDS cost is defined as

$$E_{\text{mds}} = \|HDH - H\text{dists}(X)H\|_F^2,$$

where $H = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T$. The robust embedding cost is

$$E_{\text{robust}} = \sum_{ij} |D_{ij} - [\text{dists}(X)]_{ij}| \,.$$

In both cases, D is the matrix of squared dissimilarities, and dists(X) is the matrix of squared Euclidean distances.

Figure shows the plot of E_{mds} for the two different embeddings (MDS and SDS) as a function of dimensionality. The errors are essentially identical for the two different embeddings and there is no clear elbow.

Figure shows the plot of E_{robust} for the two embeddings. Again, the error of the two embeddings is roughly the same, though SDS is slightly better in high dimensions. What is remarkable is that there is a very clear minimum occurring at $d = 8$, which is the purported intrinsic dimensionality.

From this experiment, it appears that the robust metric really is more useful than the MDS metric. What is curious is that the original embedding method does not seem to matter. On the other hand, a subgradient algorithm was used for SDS and was stopped well before it converged.

We performed an nearly identical experiment on the faces dataset taken from the ISOMAP website. The only difference was that the 6-nearest neighbors were used instead of 7. The dimensionality reported for this dataset by Hein & Audibert is slightly different than our predictions: we predict 5, they predict 3.

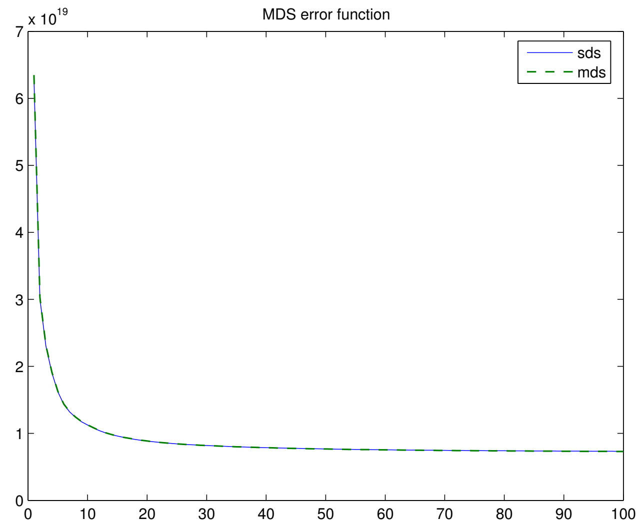

Figure 1: Ones dataset. MDS error as a function of dimensionality.

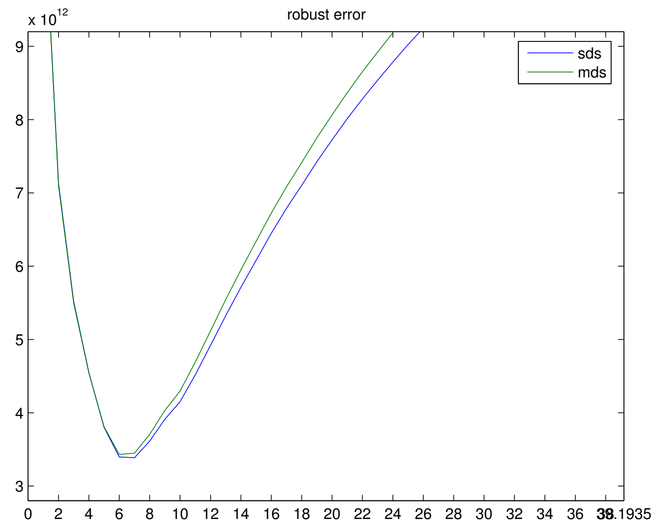

Figure 2: Ones dataset. Robust error as a function of dimensionality.

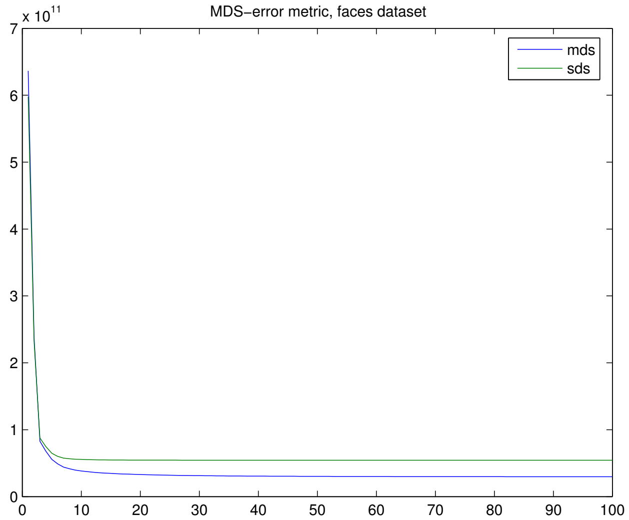

Figure 3: Faces dataset. MDS error as a function of dimensionality

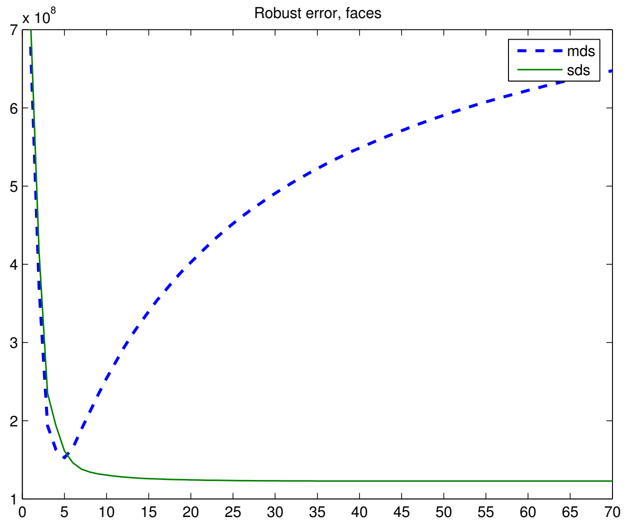

Figure 4: Faces dataset. Robust error as a function of dimensionality.