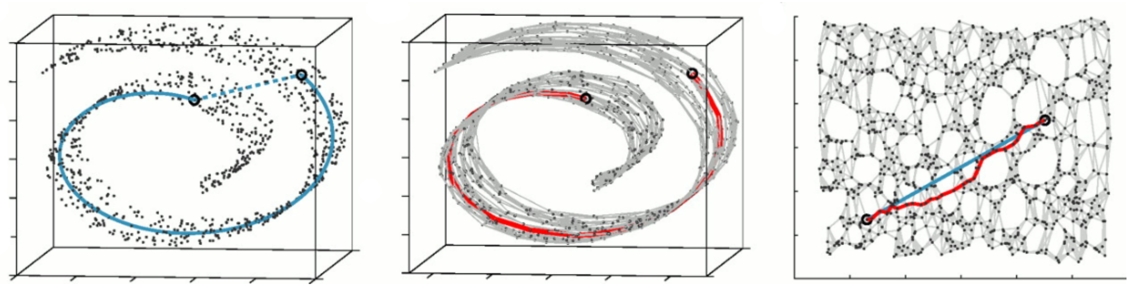
Isomap Implementation Issues

Isomap (isometric feature mapping) is a nonlinear dimensionality reduction method. It is most widely used in low-dimensional embedding methods. Isomap is a nonlinear generalization of classical multidimensional scaling (MDS). The main idea is to perform MDS, not in the input space, but in the geodesic space of the nonlinear data manifold. The geodesic distances represent the shortest paths along the curved surface of the manifold measured as if the surface were flat. This can be approximated by a sequence of short steps between neighboring sample points. Isomap then applies MDS to the geodesic rather than straight line distances to find a low-dimensional mapping that preserves these pairwise distances. The Isomap algorithm proceeds in three steps: find the neighbors of each data point in high-dimensional data space, compute the geodesic pairwise distances between all points, and embed the data via MDS to preserve these distances.

The following are the implementation issues associated with Isomaps:

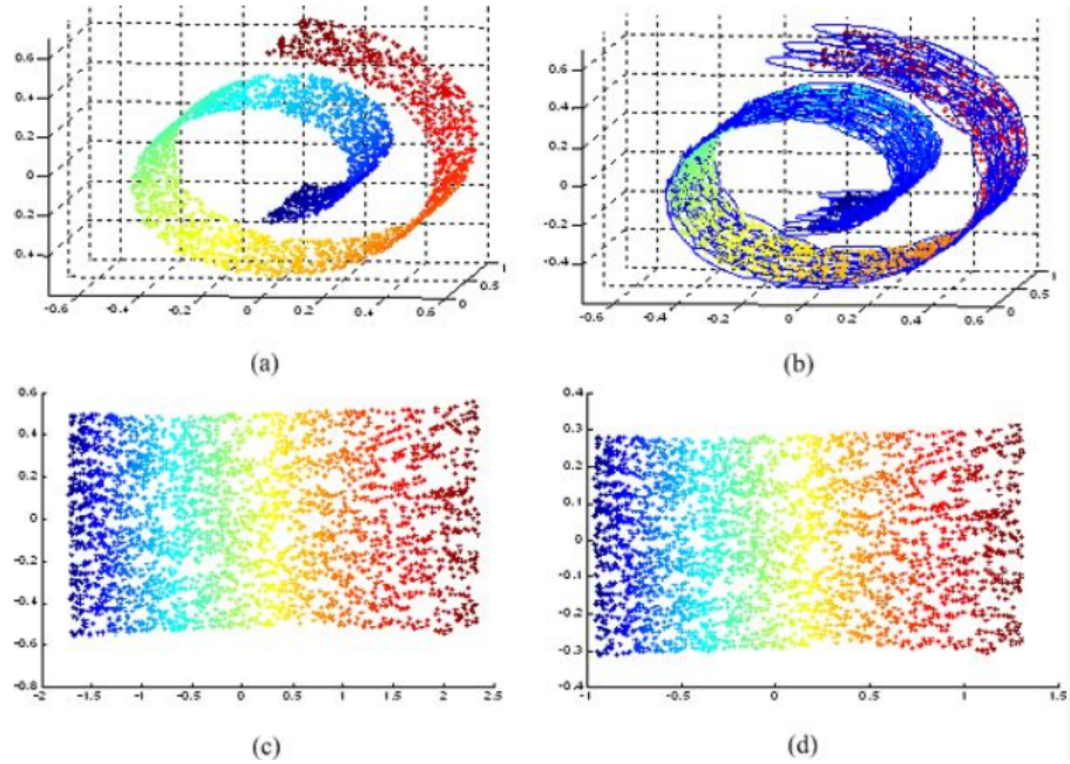
- Isomaps are computationally too expensive (dense matrix eigen-reduction) or impractical in cases where dimensionality reduction must be applied on a data stream. It is too expensive to use on any but relatively small datasets (Schoeneman, Mahapatra, Chandola, Napp, Zola).
- Isomap is a nonlinear dimensionality reduction technique whose generalization ability id very weak. The projection results in low-dimensional space of the test image cannot be easily obtained.



The figure above depicts how the swiss roll dataset which illustrates how Isomap exploits geodesic paths for dimensionality reduction. The first graph shows how the Euclidean distance between two arbitrary points in high-dimensional input space does not accurately reflect their intrinsic similarity as measured by geodesic distance along the low-dimensional manifold (length of solid curve) as seen in the second graph. Graph three shows us how the path distances are best preserved in the two-dimensional embedding recovered by Isomap.

- Isomaps are sensitive to noise and artifacts. This is not desirable since real world data is usually contaminated with noise and various artifacts due to imperfect sensors or human errors. This sensitivity is particularly relevant for emerging architectures for very low power sensing (Dadkhahi, Duarte, Marlin).
- Isomaps do not strive to preserve the local properties of data.
- Isomaps can reliably recover low-dimensional nonlinear structures in high-dimensional datasets but suffer from the problem of short-circuiting, which occurs when the neighborhood distance is larger than the distance between the folds in the manifolds (Saxena, Gupta, Mukerjee).
- The performance of the Isomap algorithm depends significantly on the number of neighbors chosen to build the nearest-neighbors graph of the given data. Using a large number of neighbors causes short circuiting in the k-NN graph, which leads to improper unfolding of the manifold.
- Isomaps perform poorly when manifold is not well sampled and

contains holes. It also fails to perform well if the manifold is non-convex.



In the swiss roll diagram above, the red and the blue points are very close in terms of Euclidean distance but very far in terms of geodesic distance

- Isomaps have no local assumptions and only need a sparse set of distance measures. This facilitates the implementation on an embedded platform.
- The Isomap algorithm is topologically insatiable. Isomaps may construct erroneous connections sin the neighborhood graphs.
- The connectivity of each data point in the neighborhood graph is defined as its nearest k Euclidean neighbors in the high-dimensional space. This step is vulnerable to ”short-circuit errors” if k is too large with respect to the manifold structure or

if noise in the data moves the points slightly off the manifold. Even a single short-circuit error can alter many entries in the geodesic distance matrix, which in turn can lead to a drastically different (and incorrect) low-dimensional embedding.

- If k is too small, the neighborhood graph may become too sparse to approximate geodesic paths accurately.
- If new training data is added to the original training data, the whole of the Isomap needs to be computed again. Adding a new datapoint has a large enough influence on the map pf neighbors and you will have to change the weights of the old data points (Hadid, Kouropteva, Pietikainen).

Schoeneman, F. Mahapatra, S. Chandola, V. Napp, N. Zola J. Error Metrics for Learning Reliable Manifolds from Streaming Data. Retrieved from: https://arxiv.org

Dadkhahi, H. Duarte, M. Marlin, B. Out-of-Sample Extension for Dimensionality Reduction of Noisy Time Series. Retrieved from: http://openscholar.cs.umass.edu

Saxena, A. Gupta, A. Mukerjee, A. Non-linear Dimensionality Reduction by Locally Linear Isomaps. Retrieved from: https://pdfs.semanticscholar.org
Hadid, A. Kouropteva, O. Pietikainen, M. Unsupervised Learning using Locally Linear Embedding: Experiments with Face Pose Analysis. Retrieved from: http://www.ee.oulu.fi