Isomap: Isometric Mapping

1 Introduction

Isomap, or Isometric Mapping, is a nonlinear dimensionality reduction technique that aims to preserve the intrinsic geometric structure of high-dimensional data by maintaining the geodesic distances between points. Unlike linear techniques such as PCA, Isomap is designed to uncover the underlying manifold structure in the data, which is particularly useful for datasets that are distributed on a nonlinear manifold. Isomap builds on classical Multidimensional Scaling (MDS) by incorporating geodesic distances instead of simple Euclidean distances, thus providing a more accurate low-dimensional representation of the data.

2 Mathematical Foundations

Isomap is grounded in several key mathematical principles that allow it to effectively map high-dimensional data to a lower-dimensional space while preserving its essential structure.

2.1 Geodesic Distance

The geodesic distance between two points on a manifold is the length of the shortest path along the surface of the manifold that connects them. This concept is crucial for Isomap because it allows the algorithm to account for the true underlying geometry of the data, which may be distorted in high-dimensional space. Unlike the Euclidean distance, which may underestimate the true distance between points on a curved surface, the geodesic distance captures the correct separation by following the manifold's curvature.

2.2 Graph Construction

To approximate the manifold's structure, Isomap represents the data as a weighted graph $G = (V, E)$, where each vertex $v_i \in V$ corresponds to a data point x_i. The edges $e_{ij} \in E$ are created between each data point and its k-nearest neighbors based on the Euclidean distance. The weight of an edge w_{ij} is the Euclidean distance between the connected points, which serves as an initial approximation of the local geodesic distance.

$$w_{ij} = \begin{cases} d(x_i, x_j) & \text{if } x_j \text{ is a neighbor of } x_i, \\ \infty & \text{otherwise.} \end{cases}$$

This step effectively captures the local neighborhood structure of the manifold.

2.3 Shortest Path Algorithms

Once the graph is constructed, Isomap uses a shortest path algorithm, such as Floyd-Warshall or Dijkstra's algorithm, to compute the shortest paths between all pairs of points. These paths represent the geodesic distances on the manifold. The idea is that by piecing together these short paths, we can approximate the global structure of the manifold.

$$d_G(x_i, x_j) = \min \left(\sum_{e_{ab} \in P_{ij}} w_{ab} \right),$$

where P_{ij} represents all possible paths between points x_i and x_j, and w_{ab} is the weight of edge e_{ab}. The computed $d_G(x_i, x_j)$ gives a more accurate measure of the true distance between points on the manifold.

2.4 Classical Multidimensional Scaling (MDS)

After computing the geodesic distances, Isomap applies classical Multidimensional Scaling (MDS) to these distances to find a low-dimensional embedding of the data. The MDS technique seeks to preserve the pairwise distances between points in the low-dimensional space, effectively capturing the global structure of the manifold. The key step in MDS involves solving an eigenvalue problem for the doubly centered distance matrix:

$$\mathbf{B} = \mathbf{H}\mathbf{D}_G\mathbf{H},$$

where $\mathbf{B}$ is the doubly centered distance matrix, and $\mathbf{H}$ is the centering matrix defined as:

$$\mathbf{H} = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^\top,$$

with $\mathbf{I}$ being the identity matrix and $\mathbf{1}$ being a vector of ones. The matrix $\mathbf{B}$ is then diagonalized to extract the top d eigenvectors, which correspond to the low-dimensional coordinates of the data points.

3 Objective Function

The goal of Isomap is to minimize the stress function, which quantifies the difference between the geodesic distances in the original high-dimensional space and the Euclidean distances in the low-dimensional embedding. The stress function is defined as:

$$\text{Stress} = \sum_{i<j} \left(d_G(x_i, x_j) - \|y_i - y_j\|\right)^2 ,$$

where y_i and y_j are the low-dimensional representations of the high-dimensional data points x_i and x_j, respectively, and $\|\cdot\|$ denotes the Euclidean distance in the low-dimensional space. By minimizing this stress function, Isomap ensures that the low-dimensional embedding preserves the manifold's original geometry as closely as possible.

4 Applications of Isomap

Isomap has been successfully applied in various fields due to its ability to reveal the intrinsic low-dimensional structure of high-dimensional data:

- **Manifold Learning:** Isomap is extensively used in manifold learning to uncover and visualize the hidden structure of complex datasets, such as those found in robotics, biology, and astronomy.

- **Image Processing:** In image processing, Isomap is used to reduce the dimensionality of image data, making it easier to analyze and process while preserving the relationships between different images.

- **Genomics:** In genomics, Isomap helps in reducing the dimensionality of gene expression data, enabling the discovery of patterns and relationships that are not easily detectable in the original high-dimensional space.

- **Speech Analysis:** Isomap has been applied to speech data to reduce its dimensionality, which improves the efficiency and accuracy of speech recognition systems.

5 Summary

Isomap is a powerful tool for nonlinear dimensionality reduction, particularly when the objective is to preserve the global structure of the data as reflected by geodesic distances. By combining graph-based distance computation with classical MDS, Isomap effectively uncovers the low-dimensional manifold structure of high-dimensional datasets, making it a valuable technique for manifold learning and various data analysis tasks.