The Isomap Algorithm

For Nonlinear Dimensionality Reduction

(Inspired in part by an AATRN talk given by Henry Adams & Universitat Tübingen Statistical ML course lectures)

2/28/2022

CompTAG Seminar

Isomap

A fundamental manifold learning technique

History: Original Paper (2000)

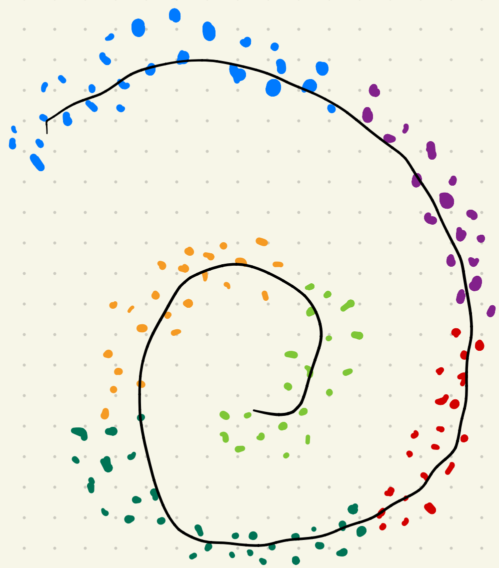
J. Tenenbaum, V. De Silva, J. Langford.
"A global geometric framework for nonlinear dimensionality reduction." in Science

→ Right at the beginning of when manifold methods were becoming fashionable in ML

→ Many, many, techniques sharing isomap philosophy were invented shortly thereafter
i.e.
(Locally Linear embedding, Laplacian eigenmaps, Hessian Eigenmaps, Diffusion Maps, Max. Variance Unfolding, etc...)

A Motivating Example:

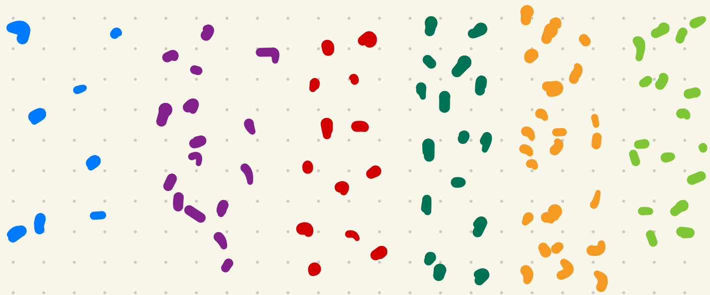
"Swiss Roll"



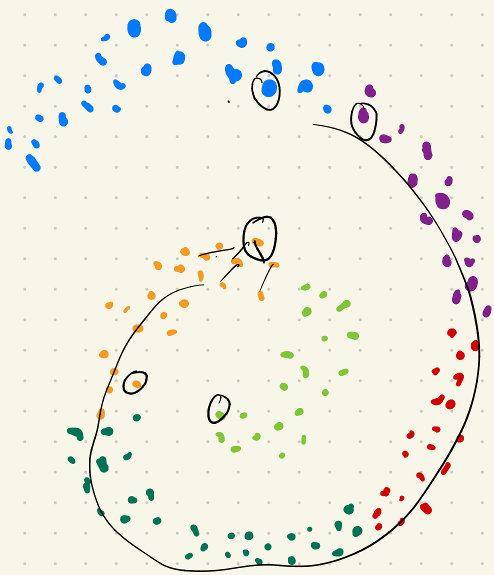
Linear Dimensionality Reduction: (PCA, for example)

- Projection onto a linear subspace
- Can, by collapsing, lose out on some of the patterns in this data!

Want: To "unroll" the "Swiss Roll"



In so doing, we attempt to move away from something like the Euclidean distance, and move toward some form of geodesic distance.



How do we construct a notion of geodesic distance from a discrete set of points?

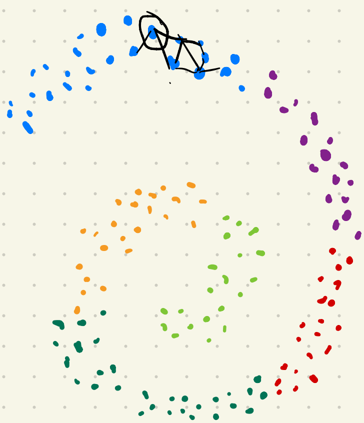
<u>Isomap:</u> Build a graph from points.

<u>Two techniques:</u>

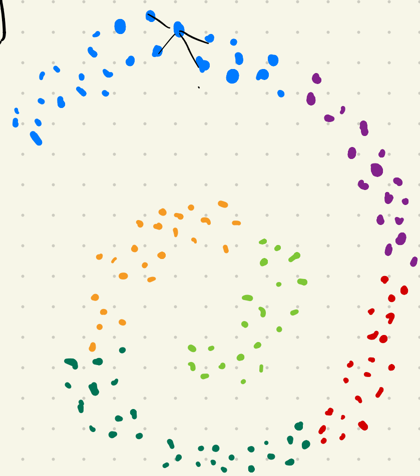
1) k-nearest neighbor, fixing k

2) add edges within a fixed metric ball

1) $k=3$

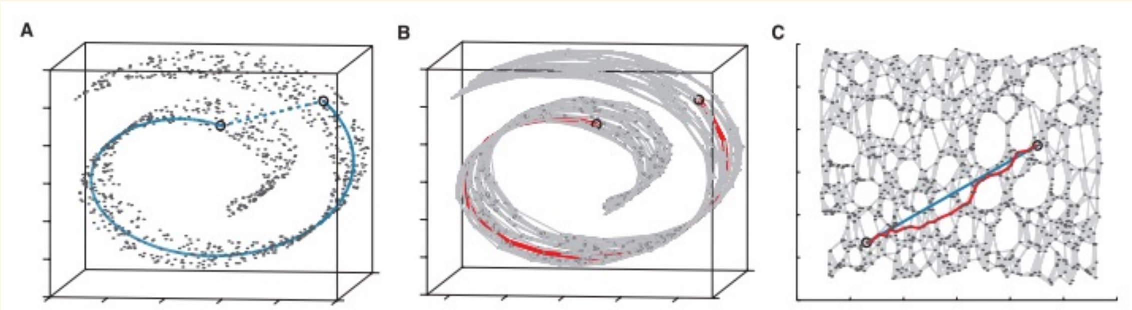


2) $\delta =$



Ok, we've constructed a network from our data.

Dimensionality Reduction is then possible by embedding our data in a lower dimensional space, s.t. the geodesic distances are as close as possible to the Euclidean distances in that space.



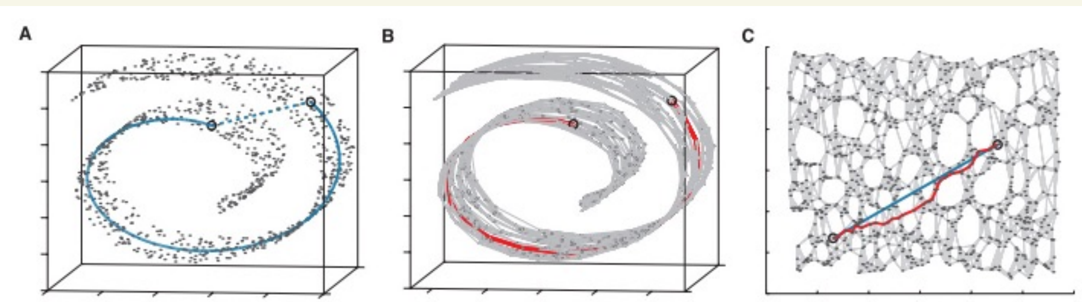
What have we done?

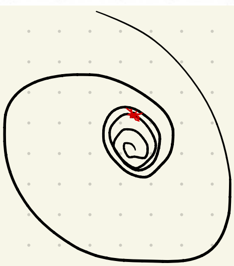
1) Began w/ abstract set of points

2.) Built k-nearest neighbor (or metric ball) inspired graph w/ edges weighted by local distances

3.) Compute geodesic distances for each pt. to every other pt., + store in $n \times n$ matrix.

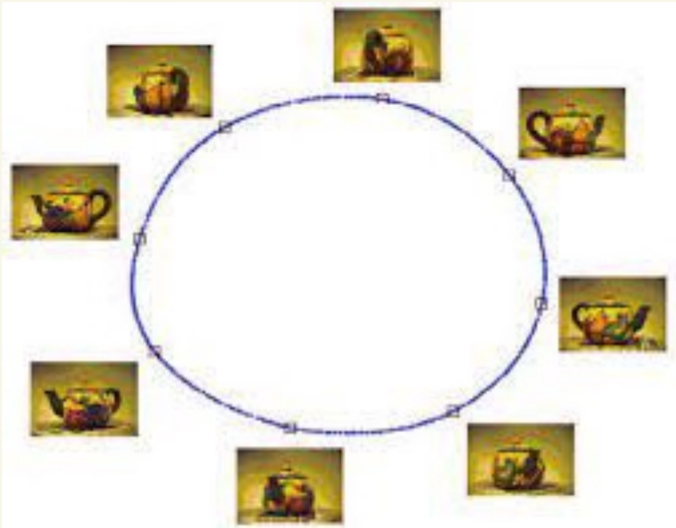
4.) Apply metric MDS with this matrix now as input. → finds embedding preserving geodesic distances





Examples:

Angles of a teapot



High dimensional, BUT should live on a 1-dim'l manifold, roughly speaking.

More Examples & Review:

Fundamentally, if we think our data should live on some k-manifold, isomap is an attempt to find and make sense of that manifold in "only" k-dimensions.