Seminar Report

Applying Semi-Supervised Locally Linear Embedding

Department of Statistics
Ludwig-Maximilians-Universität München

By Lisa Wimmer
Under the supervision of Jann Goschenhofer, Ph.D.
Munich, month day$^{\text{th}}$, 2021

Abstract

Storyline

- Goal: present SS-LLE as a local, graph-based manifold learning method incorporating prior knowledge
- Step 0: define basic mathematical concepts required to understand argumentation (plus notation)
- Step 1: introduce idea of **isometry** (most basic: MDS)
- Step 2: introduce idea of **graph-based** models
  - Achieve non-linearity
  - Common structure: build graph → derive matrix as quadratic form over graph function → derive embedding from eigenvalue problem
  - Most basic: ISOMAP (global, dense, convex)
- Step 3: introduce idea of **locality**
  - Relax global to local isometry
  - Find sparse rather than dense matrices
  - **Laplacian eigenmaps** as concept in which the others can be generalized
    - Define weighting scheme for neighborhood
    - Use Laplacian to derive matrix
    - Solve sparse eigenvalue problem
- Step 4: introduce **local linearity**
  - **LLE**
    - Obtain weights via linear reconstructions
    - Can be shown to approximate graph Laplacian (Belkin & Niyogi (2006))
  - **Hessian LLE**
    - Replace Laplacian by Hessian
- Step 5: introduce **prior knowledge**
  - **SS-LLE**
  - Improve results by pre-specifying some manifold coordinates

Contents

List of Figures

List of Tables

1 Introduction

Will surely cite a lot from LLE paper (Roweis and Saul, 2000) and SS-LLE paper (Yang et al., 2006)

- Why is dimensionality reduction desirable? Not only because it's easier to handle and visualize lower-dimensional data but because data-generating process is often truly of much lower dimension
- Our goal is to find the mapping from latent feature space embedded in the m-dimensional Euclidean space we observe to the d-dimensional space the embedding is locally homeomorphic to (unrolling the Swiss roll)
- This mapping can be constructed linearly or non-linearly (slapping the roll flat vs unrolling it), thereby defining the complexity of the manifolds we are able to learn
- Brief intuition to manifold learning with simple example (e.g., rotated letters A)
- Different methods out there (linear, non-linear, ...)

2 Mathematical Framework

2.1 "Mathematical Objects" (find proper title)

- Topological spaces
- Topological manifols
- Curves/geodesics
- Tangent spaces

2.2 Spectral Decomposition

- Eigenvalues/eigenvectors
- Spectral decomposition

3 Local Graph-Based Manifold Learning

3.1 Concept of Isometry

- Notion of distance
- Preserving distances in manifold learning
- MDS (very brief)

3.2 Graph-Based Models

3.2.1 Neighborhoods

- k-/ϵ-neighborhoods and neighborhood graphs
- Linear reconstruction and reconstruction error

3.2.2 Basics of Spectral Graph Theory

- Degree and adjacency matrices
- Laplacian operators

3.2.3 General Structure of Graph-Based Models

- Neighborhood graph
- Weight matrix
- Eigenwert problem

3.2.4 ISOMAP

- (One of the) earliest, simplest variant(s)
- MDS with geodesics

3.3 Laplacian Eigenmaps

- Notion of locality
- Laplacian eigenmaps

3.4 Locally Linear Embedding (LLE)

- Notion of local linearity
- Approximation of graph Laplacian

3.5 Hessian Locally Linear Embedding (HLLE)

- Hessian instead of Laplacian (eigenmaps)
- Hessian instead of LS fit (LLE)

4 Semi-Supervised Locally Linear Embedding (SS-LLE)

4.1 Employment of Prior Information

- Why use labels in the first place?
- How will that help?
- How do we even find prior points?
- Exact vs inexact knowledge

4.2 SS-LLE Algorithm

- What is different wrt standard LLE?

4.3 Strengths and Drawbacks of SS-LLE

Potential shortcoming: what if manifold is not well-sampled? Not a problem with synthetic data, but IRL. But probably problematic with all manifold approaches

Also: generalization to new points (w/o recomputing everything) neighborhood-preserving propositions

5 Experiment Results

5.1 Data

5.2 Experimental Design

- Implementation details
- Hyperparameters
- Evaluation criteria

5.3 Results and Discussion

6 Conclusion

Lorem ipsum

A Appendix

Lorem ipsum

B Electronic Appendix

Data, code and figures are provided in electronic form.

References

Roweis, S. T. and Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding, *Science* **290**: 2323–2326.

Yang, X., Fu, H., Zha, H. and Barlow, J. (2006). Semi-supervised nonlinear dimensionality reduction, *Proceedings of the 23rd International Conference on Machine Learning*, Pittsburgh, PA, USA.

Declaration of Authorship

I hereby declare that the report submitted is my own unaided work. All direct or indirect sources used are acknowledged as references. I am aware that the Thesis in digital form can be examined for the use of unauthorized aid and in order to determine whether the report as a whole or parts incorporated in it may be deemed as plagiarism. For the comparison of my work with existing sources I agree that it shall be entered in a database where it shall also remain after examination, to enable comparison with future Theses submitted. Further rights of reproduction and usage, however, are not granted here. This paper was not previously presented to another examination board and has not been published.