Seminar Report

Semi-Supervised Locally Linear Embedding

Department of Statistics
Ludwig-Maximilians-Universität München

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

Abstract

- 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
- Give brief intuition to manifold learning: what do we mean by topologies, manifolds, maps, charts?
- What methods are already out there?

Contents

List of Figures

List of Tables

1 Introduction

- 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
- Give brief intuition to manifold learning: what do we mean by topologies, manifolds, maps, charts?
- What methods are already out there?

2 Theoretical Framework

2.1 Basic Concepts in Manifold Learning

AS BRIEF AS POSSIBLE, but concepts and corresponding notation should be introduced properly

2.1.1 "Mathematical Objects" (find proper title)

- Topologies
- Manifolds
- Tangent spaces

2.1.2 Spectral Decomposition

- Eigenvalues/eigenvectors
- Spectral decomposition

2.1.3 Evaluation of Manifold Learning

- How can we evaluate manifold learning?
- No details, but some foundation to base evaluation of experiments on

2.2 Neighborhoods and Basics of Spectral Graph Theory

- Neighborhoods and neighborhood graphs
- Linear reconstruction and reconstruction error
- Degree and adjacency matrices

2.3 Laplacian Eigenmaps

- Laplacian-Beltrami operator
- Laplacian eigenmaps

Can be interpreted as framework for both LLE and HLLE; only to the extent where it helps to frame LLE

3 Locally Linear Embedding (LLE)

3.1 Original LLE Algorithm

3.2 Strengths and Drawbacks

3.3 Extensions to LLE

- Brief explanation of Hessian LLE
- Brief motivation of SS-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?

4.2 SS-LLE Algorithm

- What is different wrt standard LLE?

4.3 Choice of Prior Points

- How can we find prior points (if not by domain knowledge)?
- How many are considered necessary to yield tangible benefits?

5 Experiment Results

5.1 Data

5.2 Experimental Design

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.

Declaration of Authorship

I hereby declare that the Thesis 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 Thesis 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.