

# Hierarchical Principal Manifold Estimation

Robert Zielinski<sup>1</sup>, Kun Meng<sup>2</sup>, and Ani Eloyan<sup>1</sup>

<sup>1</sup>Department of Biostatistics, Brown University

<sup>2</sup>Division of Applied Mathematics, Brown University

August 9, 2023

## Abstract

## 1 Introduction

Neurodegenerative diseases such as Alzheimer’s disease (AD) and Parkinson’s disease are highly complex, chronic conditions that exhibit substantial heterogeneity between patients. To enable researchers to identify potential biomarkers of progression for these conditions, large multi-site longitudinal observational studies have been conducted that collect information from a wide range of participants over long durations across several modalities, including neuroimaging [Cite ADNI and PPMI studies]. Neuroimaging data in particular have been shown to be sensitive to developments in the early stages of disease progression, often before diagnosis occurs [Find citation from reading].

While observations from neuroimaging data are typically high-dimensional, it is commonly assumed that the data lie along a low-dimensional manifold [Relevant citations demonstrating this]. Under this assumption, manifold learning algorithms can be used to recover this underlying manifold and parameterize this high-dimensional input data in the learned low-dimensional space. Most commonly-used manifold learning algorithms, including Isomap (Tenenbaum, Silva, and Langford 2000), Locally Linear Embedding (Roweis and Saul 2000), Laplacian Eigenmaps (Belkin and Niyogi 2003), and Diffusion Maps (Coifman and Lafon 2006) construct a graph of similarities between observations, which is then used to determine the structure of the manifold. Meng and Eloyan 2021 presents an alternative principal manifold framework for manifold estimation which extends to higher dimensions the principal curve framework introduced in Hastie and Stuetzle 1989, in which principal curves are defined as one-dimensional curves that pass through the middle of a high-dimensional dataset.

In neuroimaging settings, manifold learning algorithms are most often used to estimate the underlying manifold at the population level, with vectorized image intensities being used as the input data [Citations]. Yue et al. [2016](#) used manifold learning to estimate a low-dimensional parameterization of the corpus callosum at the individual level. However, in many cases, the outcome of interest is framed as a comparison between different groups (e.g. treatment groups in a clinical trial).

## **2 Methods**

## **3 Simulations**

## **4 Applications**

## **5 Discussion**

## References

- Alexander, Monica and Leontine Alkema (2018). “Global Estimation of Neonatal Mortality Using a Bayesian Hierarchical Splines Regression Model”. In: *Demographic Research* 38, pp. 335–372. ISSN: 1435-9871. JSTOR: [26457049](#). URL: <https://www.jstor.org/stable/26457049> (visited on 07/06/2023).
- Belkin, Mikhail and Partha Niyogi (June 1, 2003). “Laplacian Eigenmaps for Dimensionality Reduction and Data Representation”. In: *Neural Computation* 15.6, pp. 1373–1396. ISSN: 0899-7667. DOI: [10.1162/089976603321780317](#). URL: <https://doi.org/10.1162/089976603321780317> (visited on 02/24/2023).
- Bhatia, Kanwal K. et al. (2012). “Hierarchical Manifold Learning”. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*. Ed. by Nicholas Ayache et al. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 512–519. ISBN: 978-3-642-33415-3. DOI: [10.1007/978-3-642-33415-3\\_63](#).
- Coifman, Ronald R. and Stéphane Lafon (July 1, 2006). “Diffusion Maps”. In: *Applied and Computational Harmonic Analysis*. Special Issue: Diffusion Maps and Wavelets 21.1, pp. 5–30. ISSN: 1063-5203. DOI: [10.1016/j.acha.2006.04.006](#). URL: <https://www.sciencedirect.com/science/article/pii/S1063520306000546> (visited on 08/08/2023).
- Deng, Jiansong et al. (July 1, 2008). “Polynomial Splines over Hierarchical T-meshes”. In: *Graphical Models* 70.4, pp. 76–86. ISSN: 1524-0703. DOI: [10.1016/j.gmod.2008.03.001](#). URL: <https://www.sciencedirect.com/science/article/pii/S1524070308000039> (visited on 07/06/2023).
- Forsey, David R. and Richard H. Bartels (Apr. 1995). “Surface Fitting with Hierarchical Splines”. In: *ACM Transactions on Graphics* 14.2, pp. 134–161. ISSN: 0730-0301, 1557-7368. DOI: [10.1145/221659.221665](#). URL: <https://dl.acm.org/doi/10.1145/221659.221665> (visited on 07/06/2023).
- Gao, Siyuan, Gal Mishne, and Dustin Scheinost (2019). “A Hierarchical Manifold Learning Framework for High-Dimensional Neuroimaging Data”. In: *Information Processing in Medical Imaging*. Ed. by Albert C. S. Chung et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 631–643. ISBN: 978-3-030-20351-1. DOI: [10.1007/978-3-030-20351-1\\_49](#).
- Gelman, Andrew, John B. Carlin, et al. (2014). *Bayesian Data Analysis*. 3rd. Texts in Statistical Science. Boca Raton: CRC Press/Taylor & Francis Group.
- Gelman, Andrew and Jennifer Hill (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Analytical Methods for Social Research. Cambridge: Cambridge University Press.
- Green, P. J. and B.W. Silverman (1994). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Monographs on Statistics and Applied Probability 58. Springer.

- Hastie, Trevor and Werner Stuetzle (June 1989). “Principal Curves”. In: *Journal of the American Statistical Association* 84.406, pp. 502–516. ISSN: 0162-1459, 1537-274X. DOI: [10.1080/01621459.1989.10478797](https://doi.org/10.1080/01621459.1989.10478797). URL: <http://www.tandfonline.com/doi/abs/10.1080/01621459.1989.10478797> (visited on 11/29/2022).
- He, Ying et al. (2006). “Manifold T-Spline”. In: *Geometric Modeling and Processing - GMP 2006*. Ed. by Myung-Soo Kim and Kenji Shimada. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 409–422. ISBN: 978-3-540-36865-6. DOI: [10.1007/11802914\\_29](https://doi.org/10.1007/11802914_29).
- Hendricks, Wallace and Roger Koenker (Mar. 1, 1992). “Hierarchical Spline Models for Conditional Quantiles and the Demand for Electricity”. In: *Journal of the American Statistical Association* 87.417, pp. 58–68. ISSN: 0162-1459. DOI: [10.1080/01621459.1992.10475175](https://doi.org/10.1080/01621459.1992.10475175). URL: <https://www.tandfonline.com/doi/abs/10.1080/01621459.1992.10475175> (visited on 07/06/2023).
- Meng, Kun and Ani Eloyan (2021). “Principal Manifold Estimation via Model Complexity Selection”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 83.2, pp. 369–394. ISSN: 1467-9868. DOI: [10.1111/rssb.12416](https://doi.org/10.1111/rssb.12416). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/rssb.12416> (visited on 11/29/2022).
- Roweis, Sam T. and Lawrence K. Saul (Dec. 22, 2000). “Nonlinear Dimensionality Reduction by Locally Linear Embedding”. In: *Science* 290.5500, pp. 2323–2326. DOI: [10.1126/science.290.5500.2323](https://doi.org/10.1126/science.290.5500.2323). URL: <https://www.science.org/doi/10.1126/science.290.5500.2323> (visited on 02/24/2023).
- Schulam, Peter and Suchi Saria (2016). “Integrative Analysis Using Coupled Latent Variable Models for Individualizing Prognoses”. In: *Journal of Machine Learning Research* 17.232, pp. 1–35. ISSN: 1533-7928. URL: <http://jmlr.org/papers/v17/15-436.html> (visited on 07/06/2023).
- Tenenbaum, Joshua B., Vin de Silva, and John C. Langford (Dec. 22, 2000). “A Global Geometric Framework for Nonlinear Dimensionality Reduction”. In: *Science* 290.5500, pp. 2319–2323. DOI: [10.1126/science.290.5500.2319](https://doi.org/10.1126/science.290.5500.2319). URL: <https://www.science.org/doi/10.1126/science.290.5500.2319> (visited on 02/24/2023).
- Xie, Z. and G.E. Farin (Jan. 2004). “Image Registration Using Hierarchical B-splines”. In: *IEEE Transactions on Visualization and Computer Graphics* 10.1, pp. 85–94. ISSN: 1941-0506. DOI: [10.1109/TVCG.2004.1260760](https://doi.org/10.1109/TVCG.2004.1260760).
- Yue, Chen et al. (July 2, 2016). “Parameterization of White Matter Manifold-Like Structures Using Principal Surfaces”. In: *Journal of the American Statistical Association* 111.515, pp. 1050–1060. ISSN: 0162-1459.

DOI: [10.1080/01621459.2016.1164050](https://doi.org/10.1080/01621459.2016.1164050). pmid: 28090127. URL: <https://doi.org/10.1080/01621459.2016.1164050> (visited on 02/24/2023).