Applying Semi-Supervised Locally Linear Embedding

Review - Extended Abstract

Presenter: Lisa Wimmer

Reviewer: Christoph Luther

January 2021

Review

The abstract on Semi-Supervised Locally Linear Embeddings (SS-LLE) is a well-structured introduction to not only the single technique but the broader context of manifold learning. The author manages to introduce the topic in a concise and accessible manner without technical details. For the purpose of this review, I will make use of the original structure of the abstract since it is coherent and easy to follow.

First, the abstract introduces manifold learning for dimensionality reduction. To this end, it highlights the key assumption that high-dimensional data often lies in a lower-dimensional manifold and stresses the local equivalence of manifolds to Euclidean spaces. Due to this local property of manifolds, neighbourhood structures in the original data can be exploited to learn such lower-dimensional manifolds. The two corresponding paragraphs explain the key ideas very well, should, however, not be expanded in the paper as the reader can be assumed to be familiar with the matter.

Second, multi-dimensional scaling (MDS) and Isomap are used to explain the notion of global dimensionality reduction techniques and to distinguish between linear and non-liner methods. Since they only serve this purpose, it would have been helpful to clarify the notion of globality, i.e. also mapping distant points in the original data to faraway embeddings and not only focusing on correctly mapping close points. The idea of approximating geodesics in Isomap, on the contrary, does not serve this purpose and may confuse the reader, as it is not clear how these geodesics come about. After introducing the notion of global dimensionality reduction techniques, local methods, which include Locally Linear Embeddings (LLE), are explained. The introduction is accessible upon first reading and very concise, yet non-technical. In my opinion, the uninformed reader could benefit from a brief explanation of how neighbourhood structures are captured in a graph (and subsequently the corresponding matrices), i.e. by constructing graphs based on the k nearest neighbours of a data point or an ϵ-radius neighbourhood. For the purpose of the paper, the paragraph can either be shortened, in order to focus solely on the abstract notion of *global* vs *local* and *linear* vs *non-linear* or extended as suggested.

The subsequent transition from Laplacian Eigenmaps to LLE and the purpose of the the weight matrix in the latter are very well explained. In order to additionally emphasize the notion of locality, it could have been explicitly stated that the weights in the linear reconstruction of points are restricted to be zero for data points not classified as neighbours of the targetted point, albeit being included in the weight matrix. Finally, the explanation of how prior information can be added to expand LLE to SS-LLE can be understood easily. On a final note, the role of Hessian LLE (H-LLE) does not become clear since it is not further explained and also not clarified whether it will be covered in the paper.

Altogether, I would suggest to keep the introduction to manifold learning and other techniques, namely MDS, Isomap and Laplacian Eigenmaps, as short as possible. As opposed to this, I would recommend to comprehensively introduce Hessian LLE as another alternative. When it comes to the implementation of methods, I would suggest to implement LLE and possibly H-LLE besides SS-LLE to be able to compare the perforamce of all methods for different data sets.