Probabilistic PCA

Roxana Shaikh

December 21, 2020

1 Objective

Principal Component Analysis (PCA) is a popular and effective technique for dimensionality reduction. It performs a linear transformation to a new coordinate system, which emphasizes the dimensions of the dataset that show the highest variances.

One of the strengths of PCA can at the same time be viewed as one of its biggest drawbacks - there are no parameters which can impact the performance of conventional PCA. This also carries the effect, that there is no way to tweak and adjust the results. Furthermore PCA is a method which maps the data in a linear way, which

The objective of the upcoming paper is to highlight the advantages of Probabilistic PCA (PPCA) in comparison to conventional PCA and show the different variations of PPCA.

2 Probabilistic PCA

There have been a few proposals towards nonlinear approaches to PCA. One possibility to circumvent this is to assume an underlying probability density, proposed by Tipping and Bishop (1997, 1999b) and also by Roweis (1998).

Generally Probabilistic PCA represents method, where the main constraint is the chosen underlying probability distribution which still offers room to captivate the dominant attributes of the observation set.

2.1 Maximum Likelyhood PCA (MLPCA)

MLPCA is a decomposition method similar to conventional PCA, however it takes into account measurement uncertainty in the decomposition process. Its general idea can be emcompassed by the assignment of a correspondent latent variable to each data point.

Making use of the EM-algorithm instead of the covariance matrix is an computationally effecient way of performing MLPCA for high dimensional data.

2.2 Bayesian PCA

One big problematic aspect of PCA as well as MLPCA is to retain the ideal number or principal components, which is no other than finding out the number of dimensions that truly matter. The Bayesian approach is an automated way of determining the appropriate dimensionality.

Bayesian PCA means formulating PCA as a Bayesian model of the statistical data. When using a Bayesian model, a latent dimensionality is automatically chosen by maximizing the approximate marginal likelihood of the model.

2.3 Mixtures of Probabilistic Principal Component Analysers (MPPCA)

Using a probabilstic model as a base for PCA leads inevitebly to the continuation of applying local probability densities to each dimension, which results in a mixture of of probabilistic principal component analysers.

The mixture of probabilistic principal component analysers allows for limited control over the free parameters by choosing the latent space dimension. However it is not nearly as restrictive as the constraints that are put up by traditional PCA, which are very generalized and thus can be misleading.

3 Conclusion

Applying an underlying probability density to PCA results in opening the door in my directions for further exploration and analysis, as now there is the possibility to apply techniques that are intended for probability densities, as an appropriate density model of the observations has been computed.

To be more specific, this means Bayesian inference methods can be applied, but also comparisons to other methods which use density estimations can be drawn.

4 References

Christopher M. Bishop. 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag, Berlin, Heidelberg. Chapter 12.2

Tipping ME, Bishop CM. Mixtures of probabilistic principal component analyzers. Neural Comput. 1999 Feb 15;11(2):443-82. doi: 10.1162/089976699300016728. PMID: 9950739.

Tipping, M. E., Bishop, C. M. (1999). Probabilistic principal component analysis. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 61(3), 611–622.

A. Ghodsi, “Dimensionality Reduction A Short Tutorial,” 2006.