



Density Estimation and Classification via Bayesian Nonparametric Learning of Affine Subspaces

[Abhishek Bhattacharya](#), [Garritt Page](#), [David Dunson](#)

(Submitted on 28 May 2011)

It is now practically the norm for data to be very high dimensional in areas such as genetics, machine vision, image analysis and many others. When analyzing such data, parametric models are often too inflexible while nonparametric procedures tend to be non-robust because of insufficient data on these high dimensional spaces. It is often the case with high-dimensional data that most of the variability tends to be along a few directions, or more generally along a much smaller dimensional submanifold of the data space. In this article, we propose a class of models that flexibly learn about this submanifold and its dimension which simultaneously performs dimension reduction. As a result, density estimation is carried out efficiently. When performing classification with a large predictor space, our approach allows the category probabilities to vary nonparametrically with a few features expressed as linear combinations of the predictors. As opposed to many black-box methods for dimensionality reduction, the proposed model is appealing in having clearly interpretable and identifiable parameters. Gibbs sampling methods are developed for posterior computation, and the methods are illustrated in simulated and real data applications.

Subjects: **Methodology (stat.ME)**

Cite as: [arXiv:1105.5737](#) [stat.ME]

(or [arXiv:1105.5737v1](#) [stat.ME] for this version)

Submission history

From: [Garritt Page](#) [[view email](#)]

[v1] Sat, 28 May 2011 20:11:25 GMT (160kb,D)

[Which authors of this paper are endorsers?](#)

Download:

- [PDF](#)
- [Other formats](#)

Current browse context:

stat.ME

[< prev](#) | [next >](#)

[new](#) | [recent](#) | [1105](#)

Change to browse by:

[stat](#)

References & Citations

- [NASA ADS](#)

Bookmark (what is this?)

