ACOUSTICS2008/1471
Bayesian inference in hierarchical non-negative matrix factorisation models of musical sounds

Ali Taylan Cemgil[a] and Tuomas Virtanen[b]
[a]University of Cambridge, Trumpington street, CB2 1PZ Cambridge, UK
[b]Tampere University of Technology, Inst. of Signal Processing, FI-33101 Tampere, Finland

There has been a surge of interest to efficient audio and music modelling using tools from statistical machine learning. One such technique, that has been particularly successful, is non-negative matrix factorisation (NMF). However, a detailed theoretical understanding of this success is missing, as well as links to other modelling strategies such as sinusoidal or harmonic models. To fill this gap, we describe NMF in a statistical framework, as a hierarchical generative model consisting of an observation and a prior component. We show that particular choices lead to standard NMF algorithms as special cases, where parameter estimation is carried out via maximum likelihood. Starting from this view, we develop extensions that facilitate more realistic acoustic modelling (such as spectral smoothness or harmonicity of natural sounds) and alternative inference techniques via Gibbs sampling and variational Bayes, which allow us to do principled comparisons between alternative models via Bayesian model selection. Our novel construction, where we make use of Markov chains of Gamma random variables, retains conjugacy and enables us to develop models that fit better to real data while retaining attractive features of standard NMF such as fast convergence and easy implementation. We illustrate our approach on polyphonic pitch estimation.