Probabilistic models for structured sparsity

Sparsity has become an increasingly popular choice of regularization in machine learning and statistics. The sparsity assumption for a matrix $X$ means that most of the entries in $X$ are equal to exactly zero. Structured sparsity is a generalization of sparsity and assumes that the set of locations of the non-zero coefficients in $X$ contains structure that can be exploited. This thesis deals with probabilistic models for structured sparsity for regularization of ill-posed problems. The aim of the thesis is two-fold; to construct sparsity promoting prior distributions for structured sparsity and to derive efficient inference algorithms for these distributions. The work explores a class of models that uses Gaussian processes (Rasmussen and Williams, 2006) as a latent representation of the structure of sparsity patterns. This representation allows prior knowledge of the structure of the sparsity patterns to be encoded using generic covariance functions through the Gaussian process. This thesis focuses on two specific instances of ill-posed problems: linear inverse problems and time-varying covariance estimation. The first part of the thesis deals with probabilistic methods for finding structured sparse solutions to linear inverse problems. In this part, the sparsity promoting prior known as the spike-and-slab prior (Mitchell and Beauchamp, 1988) is generalized to the structured sparsity setting. An expectation propagation algorithm is derived for approximate posterior inference. The proposed model and the associated inference algorithm are studied and evaluated using a set of numerical experiments, which include phase transition experiments, compressed sensing, phoneme classification and electroencephalography (EEG) source localization. The second part of the thesis deals with the problem of time-varying covariance estimation. A hierarchical model for a set of non-stationary time series with time-varying covariance matrices is proposed. The model is tailored to address the problem of dynamic functional connectivity in neuroimaging and it assumes that the instantaneous covariance matrix of each time series is decomposed into a non-negative linear combination of elements from a dictionary of shared covariance matrix components. A variational Bayes algorithm is derived for approximate posterior inference. The proposed model is validated using a functional magnetic resonance imaging (fMRI) dataset.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems
Contributors: Andersen, M. R.
Number of pages: 181
Publication date: 2017

Publication information
Publisher: Technical University of Denmark (DTU)
Original language: English
Electronic versions:
phd452_Andersen_MR_002_.pdf