10th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2017)
16-18 December 2017, Senate House, University of London, UK
ERCIM WG CMStatistics
Birkbeck UNIVERSITY OF LONDON
UNIVERSITY OF LONDON

Sessions on Graphical Markov Models at CMStat 2017

Session EO298 — ***Room:*** **MAL B18**
Graphical Markov models I — *Saturday 16.12.2017* 11:25 - 13:05
Chair: Giovanni Marchetti — *Organizer:* Giovanni Marchetti

EO0911: A. Roverato, L. La Rocca
Undirected, indirected and regression graph models for categorical data in a common framework
EO0436: H. Massam
Precision matrix estimation in large coloured graphical Gaussian models
replaced by Monia Lupparelli "Distortion of effects in discrete graphical models"
EO1291: R. Evans
Model selection and local geometry
EO1306: S. Massa
Learning stable graphical models

Session EO019 — ***Room:*** **MAL B18**
Graphical Markov models II — *Saturday 16.12.2017* 14:35 - 16:15
Chair: Nanny Wermuth — *Organizer:* Nanny Wermuth

EO1243: G. Marchetti
Identical maximum likelihood estimates for Gaussian and Ising models defined by a chordless cycle
EO0635: C. Tarantola, M. Lupparelli, I. Ntzoufras
Probability based independence sampler for Bayesian quantitative learning in graphical log-linear marginal models
EO1281: C. Uhler, Y. Wang, L. Solus, K. Yang
Permutation-based causal inference algorithms with interventions
EO0894: P. Zwiernik
Testing total positivity in Gaussian graphical models

Session EO740 — ***Room:*** **MAL B18**
Graphical Markov models III — *Sunday 17.12.2017* 14:25 - 16:05
Chair: Elena Stanghellini — *Organizer:* Elena Stanghellini

EO0442: G. Letac, J. Wesolowski
Graphical models with prescribed negative covariances and unknown variances
replaced by Kayvan Sadeghi "On finite exchangeability and conditional Independence"
EO1288: S. Lauritzen
A generic algorithm for estimation in undirected graphical models
EO1063: J. Mooij
Markov and other properties of cyclic structural causal models
EO1195: A. Gottard
Graphical models based on trees

Session EO734 — ***Room:*** **MAL B18**
Graphical Markov models IV — *Sunday 17.12.2017* 16:35 - 18:15
Chair: TBA — *Organizer:* Kayvan Sadeghi

EO0848: D. Rothenhaeusler, P. Buehlmann, J. Ernest
Causal inference in partially linear structural equation models with Gaussian noise
EO0793: M. Scutari
Bayesian Dirichlet Bayesian network scores and the maximum entropy principle
EO1276: N. Wermuth
On the attractive properties for estimating and generating distributions of special palindromic Ising models
EO0705: V. Didelez
Identification with graphical models for time-dependent data

Abstracts

Session E0298: Graphical Markov models I
Organizer: Giovanni Marchetti

Title: Undirected, indirected and regression graph models for categorical data in a common framework

Authors:

Alberto Roverato - Universty of Bologna (Italy) [presenting]
Luca La Rocca - University of Modena and Reggio Emilia (Italy)

Abstract: The problem of specifying a suitable parameterization for graphical models for categorical data is considered. We focus on three of the most relevant families of graphical models, that is, undirected, bidirected and regression graph models. In this respect, we first give some general properties concerning conditional independence and Moebius inversion. Next, we exploit these basic results to provide a unified approach to the parameterization of the three classes of models. The parameterizations are derived by applying, in the three cases, the same Moebius inversion formula to obtain a log-linear expansion of certain probabilities. In the undirected case this procedure leads to the usual corner-constrained parameterization of the class of log-linear models for contingency tables. This is the standard parameterization of undirected graph models, and we show that some well-known properties of this parameterization, such as the connection between vanishing terms and independence relationships, as well as the capability of defining context specific independencies, follow directly from the constructing procedure. In this way, the former properties automatically hold true also for other parameterizations based on the same constructing procedure, and we exploit this feature to present the theory of the three classes of models in a common framework.

Title: Distortion of effects in discrete graphical models

Authors:

Monia Lupparelli - University of Bologna (Italy) [presenting]

Abstract:
A regression framework for binary data is discussed with regression coefficients given by log-linear relative risks. A relative risk formula is derived to define the relationship between marginal and conditional relative risks. An interpretation and some applications are provided.

Title: Model selection and local geometry

Authors:

Robin Evans - University of Oxford (United Kingdom) [presenting]

Abstract: Model selection is a task of fundamental importance in statistics, and advances in high-dimensional model selection have been one of the major areas of progress over the past 20 years. Examples include covariate selection in linear regression, and models based on patterns of zeros in the inverse covariance matrix. Much of this progress has been due to penalized methods such as the lasso, and efficient methods for solving the relevant convex optimization problems. However in other classes, such as directed graphical models, correct model selection is provably hard. We give a geometric explanation for why standard convex penalized methods cannot be adapted to directed graphs, based on the local geometry of the different models at points of intersection. These results also show that it is `statistically' hard to learn these models, and that much larger samples will typically be needed for moderate effect sizes. This has implications for other types of graphical model selection, and especially for causal models, as well as time series models. We provide some relevant heuristics that give insights into the feasibility of model selection in various classes of graphical model, including ancestral graph models, LWF chain graph models, and nested Markov models.

Title: Learning stable graphical models

Authors:

Sofia Massa - University of Oxford (United Kingdom) [presenting]

Abstract: Ongoing advances in model selection techniques for graphical models are trying to capture the structure of complex, high-dimensional datasets. Sparsity is usually invoked and techniques based on regularization, cross-validation, resampling and shrinkage estimation are becoming quite standard. One practical challenge in many applied contexts is how to assess the stability of different dependency structures and how to report the uncertainty associated with them. We will look at possible stability and uncertainty measures for undirected and chain graphs models.

Session EO019: Graphical Markov models II
Organizer: Nanny Wermuth

Title: Identical maximum likelihood estimates for Gaussian and Ising models defined by a chordless cycle
Authors:
Giovanni Marchetti - University of Florence (Italy) [presenting]
Abstract: Undirected graphical models defined by a chordless cycle require in general an iterative fitting procedures to get maximum likelihood estimates. For Gaussian models, the canonical parameters are the concentrations, that is the off-diagonal element in the inverse covariance matrix, while for Ising models, are the log-linear, two-factor interactions. However, we show conditions under which, if the canonical parameters are transformed to partial correlations, the two different likelihood functions, one for the continuous and the other for the binary variables, give the same maximum likelihood estimates provided the relevant starting correlation matrices coincide and have a closed form.

Title: Probability based independence sampler for Bayesian quantitative learning in graphical log-linear marginal models
Authors:
Claudia Tarantola - University of Pavia (Italy) [presenting]
Monia Lupparelli - University of Bologna (Italy)
Ioannis Ntzoufras - AUEB (Greece)
Abstract: A fully automatic and efficient MCMC strategy is presented for quantitative learning for graphical log-linear marginal models. While the prior is expressed in terms of the marginal log-linear parameters, we build an MCMC algorithm which employs a proposal on the probability parameter space. The corresponding proposal on the marginal log-linear interactions is obtained via parameter transformations. By this strategy, we achieve to move within the desired target space. At each step we directly work with well-defined probability distributions. Moreover, we can exploit a conditional conjugate setup to build an efficient proposal on probability parameters. The proposed methodology is illustrated using a popular four-way dataset.

Title: Permutation-based causal inference algorithms with interventions
Authors:
Yuhao Wang - MIT (United States)
Liam Solus - KTH Royal Institute of Technology (Sweden)
Karren Yang - MIT (United States)
Caroline Uhler - Massachusetts Institute of Technology (United States) [presenting]
Abstract: A recent break-through in genomics makes it possible to perform perturbation experiments at a very large scale. In order to learn gene regulatory networks from the resulting data, efficient and reliable causal inference algorithms are needed that can make use of both, observational and interventional data. We will present the first provably consistent such algorithm. It is a hybrid approach that uses conditional independence relations in a score-based method. Hence, this algorithm is non-parametric, which makes it useful for analyzing inherently non-Gaussian gene expression data. We will end by analyzing its performance on simulated data, protein signaling data, and single-cell gene expression data.

Title: Maximum likelihood estimation of the latent class model through model boundary decomposition
Authors:
Piotr Zwiernik - Universitat Pompeu Fabra (Spain) [presenting]
Abstract: The Expectation-Maximization (EM) algorithm is routinely used for the maximum likelihood estimation in the latent class analysis. However, the EM algorithm comes with no guarantees of reaching the global optimum. We study the geometry of the latent class model in order to understand the behavior of the maximum likelihood estimator. In particular, we characterize the boundary stratification of the binary latent class model with a binary hidden variable. For small models, such as for three binary observed variables, we show that this stratification allows exact computation of the maximum likelihood estimator. In this case we use simulations to study the performance of the EM algorithm.

Session EO740: Graphical Markov models III
Organizer: Elena Stanghellini

Title: On finite exchangeability and conditional independence
Authors:
Kayvan Sadeghi – University of Cambridge (United Kingdom)[presenting]
Abstract:
The independence structure of finitely exchangeable distributions over random vectors and random networks are studied. In particular, necessary and sufficient conditions for an exchangeable vector are given so that its elements are completely independent or completely dependent. These results and conditions are generalized to exchangeable random networks.

Title: A generic algorithm for estimation in undirected graphical models
Authors:
Steffen Lauritzen - University of Copenhagen (Denmark) [presenting]
Abstract: A generic and globally convergent algorithm is presented for estimation in undirected graphical models with only pairwise interaction terms, which can include penalty functions of lasso type, total positivity restrictions, and specific restrictions on the graph type. The algorithm is essentially a variant of Iterative Proportional Scaling and contains the latter as a special instance.

Title: Markov and other properties of cyclic structural causal models
Authors:
Joris Mooij - University of Amsterdam (Netherlands) [presenting]
Abstract: Structural Causal Models (SCMs), also known as (Non-Parametric) Structural Equation Models (NP-SEMs), are widely used for causal modeling purposes. One of their advantages over other representations such as causal Bayesian networks is that SCMs allow for cycles (causal feedback loops). The presence of cycles adds many complexities that are absent in the acyclic setting, especially for nonlinear models. We will discuss recent advances in the theory of cyclic Structural Causal Models. We will discuss how they can be marginalized to describe a subsystem of interest, explain how SCMs can model the equilibrium states of ordinary differential equation models and how these equilibrium states change under perturbations, and we will present recent results on the Markov properties of cyclic SCMs that provides the corner stone for novel cyclic causal discovery algorithms.

Title: Graphical models based on trees
Authors:
Anna Gottard - University of Firenze (Italy) [presenting]
Abstract: Graphical models have been in a wide range of problems to characterise the conditional independence structure among random variables. Particularly interesting are applications in genomics and omics science. A better understanding of the association among gene/protein/metabolite molecular signatures potentially offers new insights for complex diseases. With continuous random variable, most of this research traditionally focuses on Gaussian graphical models, assuming linear relationships. However, the assumptions of multivariate Gaussianity and linearity in the dependence structure are often evidently erroneous. Recent literature explores graphical models with non-linear relations. We investigate pairwise graphical models on a set of random variables, with distributions in which dependence occurs through the expected value. We study the utility of tree based models to detect interactions and non-linearities in these distributions and compare different algorithm for searching a tree or a sum of trees. A particular case of quasi-linear systems is analysed, embedding linear and nonlinear effects.

Session EO734: Graphical Markov models IV
Organizer: Kayvan Sadeghi

Title: Causal inference in partially linear structural equation models with Gaussian noise
Authors:
Dominik Rothenhaeusler - ETH Zurich (Switzerland) [presenting]
Peter Buehlmann - ETH Zurich (Switzerland)
Jan Ernest - ETH Zurich (Switzerland)
Abstract: The identifiability and estimation of partially linear additive structural equation models with Gaussian noise (PLSEMs) is considered. Existing identifiability results in the framework of additive SEMs with Gaussian noise are limited to linear and nonlinear SEMs, which can be considered as special cases of PLSEMs with vanishing non-parametric or parametric part, respectively. We close the wide gap between these two special cases by providing a comprehensive theory of the identifiability of PLSEMs by means of (A) a graphical, (B) a transformational, (C) a functional and (D) a causal ordering characterization of PLSEMs that generate a given distribution P. In particular, the characterizations (C) and (D) answer the fundamental question to which extent nonlinear functions in additive SEMs with Gaussian noise restrict the set of potential causal models and hence influence the identifiability.

Title: Bayesian Dirichlet Bayesian network scores and the maximum entropy principle
Authors:
Marco Scutari - University of Oxford (United Kingdom) [presenting]
Abstract: A classic approach for learning Bayesian networks from data is to select the maximum a posteriori (MAP) network. In the case of discrete Bayesian networks, the MAP network is selected by maximising one of several possible Bayesian Dirichlet (BD) scores, the most famous of which is the Bayesian Dirichlet equivalent uniform (BDeu) score. Since the number of possible networks to choose from grows more than exponentially in the number of variables, the uniform prior associated with BDeu makes structure learning computationally feasible, and does not require the elicitation of prior knowledge from experts. We will discuss the impact of this uniform prior on structure learning from an information theoretic perspective, showing how BDeu may violate the maximum entropy principle when applied to sparse data. On the other hand, a previous BDs score arises from a piece-wise prior and it does not appear to violate the maximum entropy principle, even though it is asymptotically equivalent to BDeu.

Title: On the attractive properties for estimating and generating distributions of special palindromic Ising models
Authors:
Nanny Wermuth - Chalmers University of Technology (Sweden) [presenting]
Abstract: Palindromic Ising models are binary quadratic exponential distributions for marginally symmetric binary variables. Their joint probabilities have an additional, special symmetry since these probabilities listed in any lexicographic order remain unchanged after fully reversing its listed elements. The term palindromic stems from linguistics, where it was coined for sequences of characters which read forward and backward in the same way, like in the sentence `step on no pets'. Such binary distributions have been used first in statistical physics for special types of ferromagnetism, called there zero-mean Ising models. We concentrate here on attractive properties for estimating and generating their distributions when all nodes in their undirected, simple and finite graphs can be ordered and interpreted to have at most two parent nodes.

Title: Identification with graphical models for time-dependent data
Authors:
Vanessa Didelez - Leibniz Institute for Prevention Research and Epidemiology - BIPS, University of Bremen (Germany) [presenting]
Abstract: Time-dependent data, e.g. longitudinal data or event histories, form the basis of many investigations. They are typically concerned with the effects of early exposures or sequential treatments on later / repeated outcomes. Some of the issues encountered in the analyses of time-dependent data include time-varying confounding, irregular observation times, drop-out and censoring. These are problems as they may render the target parameters of interest unidentifiable, e.g. due to non-ignorable drop-out. Different types of graphical models for time-dependent data will be reviewed. We then show how these can be used to characterise situations where target parameters are identified from the available data. A notion central to this characterisation is that of stability. It essentially demands that certain aspects of the underlying joint distribution be equal across regimes of interest, e.g. in the observational regime with irregular observation times and a (possibly hypothetical) experimental regime where a fixed schedule is enforced. We will illustrate how identifiability can be based on this notion of stability in very different contexts, e.g. the identification of causal effects in survival data or the ignorability of the timing of observations in longitudinal studies.