7th International Conference on Computational and Methodological Statistics (ERCIM 2014), 6-8 December 2014, University of Pisa, Italy:

~~~~~~

Three sessions on GRAPHICAL MARKOV MODELS, Dec 7 or 8, 2014

~~~~~~

First formulations of graphical Markov models started about 40 years ago, vigorous development is ongoing. These multivariate statistical models combine two simple but most powerful notions: data generating processes in sequences of single and of joint responses and conditional independences captured by graphs. The models build on early work at the beginning of the last century by the geneticist S. Wright and the probabilist A. Markov.

One primary objective is to uncover graphical representations that lead to an understanding of data generating processes. Such processes are no longer restricted to linear relations but they contain as special cases: linear dependences, subclasses of structural equations for longitudinal studies and models for sequences of interventions. In particular, response variables may be vector variables that contain discrete or continuous components or both types.

Theory of graphical Markov models: recent developments (ES100)
Organizers: Giovanni Marchetti (University of Florence) and Nanny Wermuth (Chalmers University of Technology)
Speakers:
Kayvan Sadeghi (Carnegie Mellon University), <kayvans@andrew.cmu.edu>
Robin Evans (University of Oxford), <robin.evans@stats.ox.ac.uk>
Piotr Zwiernik (University of Berkeley), <piotr.zwiernik@gmail.com>
Monia Lupparelli (University of Bologna), <monia.lupparelli@unibo.it>

Latent variables and feedback in graphical Markov models (ES114)
Organizer: Mathias Drton (University of Washington)
Speakers:
Michael Eichler (Maastricht University), <m.eichler@maastrichtuniversity.nl>
Marloes Maathuis (ETH Zuerich), <maathuis@stat.math.ethz.ch>
Joris M. Mooij (University of Amsterdam), <j.m.mooij@uva.nl>
Elena Stanghellini (University of Perugia), <elena.stanghellini@stat.unipg.it>

Recent applications of graphical Markov models (ES57)
Organizers: Nanny Wermuth (Chalmers University of Technology, Gothenburg) and Giovanni Marchetti (University of Florence)
Speakers:
Jochen Hardt (University of Mainz), <jochen.hardt@gmx.de>
Renáta Németh (Eötvös Loránd University, Budapest), <nemethr@tatk.elte.hu>
Fabrizia Mealli (University of Florence), <mealli@disia.unifi.it>
Joe Whittaker (University of Lancaster), <Joe.Whittaker@lancaster.ac.uk>

Session ES100: **The theory of graphical Markov models: recent developments**

Speaker: **Kayvan Sadeghi**, Carnegie Mellon University
Title: **Unifying graphs and separation criteria in graphical Markov models**
Abstract:
In this talk we unify all known graphs and separation criteria used in graphical models under the definition of mixed chain graphs (MCGs) with "the separation criterion".

We provide several pairwise Markov properties that are equivalent under the compositional graphoid axioms. In addition, we show that these pairwise properties are equivalent to the global Markov property associated with the separation criterion for compositional graphoids. We show how the results collapse to the known Markov properties and to theories for subclasses of MCGs.

Speaker: **Robin Evans,** University of Oxford
Title: **Causal models with hidden variables**
Abstract:
Directed acyclic graph models (DAG models, also called Bayesian networks) are widely used in the context of causal inference, since their structure admits to a calculus of intervention. However these models are not closed under marginalization, in the sense that they cannot faithfully represent the constraints imposed on margins of distributions which jointly follow a DAG. Other classes of graphs have been introduced to deal with this, including ancestral graphs and ADMGs, but we show that these are not sufficiently rich to capture the relevant models.

We present mDAGs, a new class of graphical model which is appropriate for representing causal models when some of the variables are unobserved. Results on the Markov equivalence of these models show that when interpreted causally, mDAGs are the minimal class of graphs which can be sensibly used. Understanding such equivalences is critical for the use of causal structure learning methods, and we elucidate the state of the art.

Speaker: **Piotr Zwiernik,** University of Berkeley
Coauthor: Caroline Uhler**,** Institute of Science and Technology, Klosterneuburg
Title: **Linear covariance models and their likelihood geometry**
Abstract:
Gaussian models with linear constraints on the inverse of the covariance matrix are well understood. Since the likelihood function is concave, its maximization is straightforward. On the other hand models with linear constraints on the covariance matrix are more complicated and little is known about the behavior of the likelihood function. It has been observed that in certain small cases these models behave well unless a model is badly misspecified.

We shed some more light on the geometry of the likelihood function of general linear covariance models, showing that with high probability standard numerical maximization procedures will easily reach the global maximum. As an example we consider Brownian motion tree models.

Speaker: **Monia Lupparelli**, University of Bologna
Coauthor: Alberto Roverato, University of Bologna
Title: **Log-mean linear link function for discrete regression graph models**
Abstract:
Regression graphs represent a valuable tool for investigating the independence structure in multivariate data involving variables which can be partitioned in groups, called blocks. Variables in each block are considered on equal standing and blocks are partially ordered depending on time or subject-matter considerations. Then, these models can be seen as sequences of univariate or multivariate regressions; for a comprehensive review see Wermuth and Sadeghi (2011). Parameterizations of discrete regression graphs have been studied by Drton (2009), Rudas et al. (2010) and Marchetti and Lupparelli (2011). In particular, the latter adopts the multivariate logistic link function for modelling the sequences of regressions where the regression coefficients are marginal log-linear interactions.

Inspired on the parameterization recently developed by Roverato et al. (2013), we propose an alternative approach based on the log-mean linear link function for regression modelling where the regression coefficients are log-linear combinations of relative risks. Firstly, we show that discrete regression graph models can be simply specified by setting zero log-mean linear regression coefficients. Secondly, we compare this novel approach with other existent approaches; in particular we illustrate that, depending on the chosen parameterization, different sub-models of interest can be specified by means of further zero non-independence constraints.

<u>Session ES114: **Latent variables and feedback in graphical Markov models**</u>

Speaker: **Michael Eichler**, Maastricht University
Title: **Trek Separation and Latent Variable Models for Multivariate Time Series**
Abstract:
In systems that are affected by latent variables conditional independences are often insufficient for inference about the structure of the underlying system. One common example is a system in which four observed variables X1, X2, X3, and X4 are conditionally independent given a fifth unobserved variable Y. While there are no conditional independences among the observed variables, they must satisfy the so-called tetrad constraints. Recently, Sullivant et al. (2010) discussed such additional non-Markovian constraints and provided a characterisation in terms of low-rank conditions on submatrices of the covariance matrix. Graphically these general constraints can be identified by a new separation concept called trek separation.

In this talk, we discuss the extension of the results to the multivariate time series case. Because of the commonly present serial correlation, the results are not directly applicable. For instance, the above tetrad constraints do not hold if the variables X1, X2, X3 X4, Y (as time series) have non-zero auto-correlation. Graphically, this corresponds to the fact that any instances of the observed variables cannot be separated by a single instance of Y. As an alternative, we consider mixed graphs in which each node corresponds to a complete time series. Such graphical descriptions for time series have been considered for instance by Dahlhaus (2000) and Eichler (2007).

We show that trek separation in such graphs corresponds to low-rank conditions on the spectral matrix of the process. In particular, we obtain a spectral version of the tetrad constraints in terms of spectral coherences. We discuss tests for vanishing tetrad constraints in the frequency domain based on asymptotic results and on bootstrap techniques.

Speaker: **Marloes Maathuis**, ETH Zürich
Title: **High-dimensional causal inference with latent variables**
Abstract:
We consider the problem of learning causal information between random variables in directed acyclic graphs when allowing arbitrarily many latent and selection variables. The FCI (Fast Causal Inference) algorithm has been explicitly designed to infer conditional independence and causal information in such settings. However, FCI is computationally infeasible for large graphs.

We therefore discuss the new RFCI algorithm, which is much faster than FCI. We show some theoretical results on the algorithms, discuss issues with order-dependence, and also compare the algorithms in simulation studies. Finally, we briefly discuss causal reasoning in the presence of latent variables, using a generalized backdoor criterion.

Speaker**: Joris Mooij**, University of Amsterdam
Coauthors: Dominik Jantzen and Bernhard Schölkopf, MPI Intelligent Systems Tübingen
Title: **How structural equation models can arise from dynamical systems**
Abstract:
Ordinary differential equations (ODEs) are a very popular and extremely useful mathematical modeling tool in many applied sciences (e.g., biology, chemistry, physics, etc.). ODEs are usually not thought of as causal models. On the other hand, structural equation models (SEMs), a different mathematical modeling framework mostly applied in the social and economical sciences, are usually interpreted in a causal way. I will show that these apparently different modeling frameworks are actually quite closely related.

The main result is that under certain conditions, equilibrium distributions of ODE systems can be directly mapped onto structural equation models, preserving the right semantics under interventions. This result sheds more light on the nature of SEMs, in particular in cases where causal feedback is present. It also shows that SEMs can be used as an alternative to ODEs when time series data is absent. I will discuss some issues that arise when trying to extend this framework to allow for with latent variables.

Speaker: **Elena Stanghellini**, University of Perugia
Title: **On identifiability of causal effects in Bayesian networks**
Abstract:
Several authors proved the do-calculus to be a complete algorithm for determining, in a non-parametric Bayesian network (BN), which causal effects are identifiable. The theory does not address identifiability when assumptions about the nature of the hidden variables are made. We consider all possible BN's with five discrete variables, one of which is a direct confounder with state space of size r, where r is finite.

We establish the value k such that the map from the parameters of the joint distribution to the marginal distribution over the observable variables is generically k-to-one. When k is finite, the parameters are generically identifiable. In particular, when k=r!, i.e. the only source of non-identifiability is due to swapping of the labels of the confounder, we show that the causal effect is uniquely recovered by application of the do-calculation formula. One may be tempted to conclude that, when k is finite, causal effects can always be uniquely recovered from application of the do-calculus. However, when $k>r$, instances of models are presented with generically identifiable parameters leading to conflicting results from application of the do-calculus. A Theorem on parameter equivalent models is also presented, and an example coming from applied research is discussed.

Session ES57: **Recent applications of graphical Markov models**

Speaker: **Jochen Hardt**, University of Mainz
Title: **Long term effects of emotional parentification**
Abstract:
Parentification means that a child takes care of a parent - instead of vice versa as it should be. We expected that parentified children would develop deficits in the perception of their own needs in order to survive emotionally in the family, that is they should be prone to somatization.

The present analysis is based on two internet surveys of about 500 subjects each. Via e-mail, respondents were asked to fill out a set of questionnaires. Two indicators of possible somatization were chosen as primary response, vegetative symptoms and pain. A third primary response, depression, served as a control variable. Sequences of regressions were performed to analyze the data.

Contrary to our expectations, depressive symptoms were highly associated to maternal and paternal parentification. Pain was associated with maternal parentification, vegetative symptoms only with paternal parentification.
The hypothesis of a deficit in self-perception had to be rejected. The associations of parentification to depression are as strong as those to vegetative symptoms and pain.

Speaker: **Renáta Németh**, Eötvös Loránd University of Budapest
Coauthor: Tamas Rudas, Eötvös Loránd University of Budapest
Title: **Discrete graphical models in social mobility research**
Abstract:
Graphical models are well suited to model direct and indirect associations that are of central importance in many problems of sociology. Such relevance is apparent in research on social mobility. Graphical models are discussed in the framework of marginal models, which provides a basis for a unified view.

The marginal modeling framework relies on parameters that capture aspects of associations among the variables that are relevant for the graph and, depending on the substantive problem at hand, may lead to a deeper insight than other approaches. A general version of path models for categorical data is also introduced. These models are applied to the social status attainment process in the USA, Hungary and Czechoslovakia at the end of the last century, and shows that policies in the latter socialist countries to prevent status inheritance had little success.

Speaker: **Fabrizia Mealli,** University of Florence
Coauthors: Barbara Pacini, University of Pisa, Elena Stanghellini, University of Perugia
Title: **Identification of principal causal effects using additional outcomes in concentration graphs**
Abstract:
Unless strong assumptions are made, nonparametric identification of principal causal effects in causal studies can only be partial and bounds (or sets) for the causal effects are established. In the presence of a secondary outcome, recent results exist to sharpen the bounds that exploit conditional independence assumptions. More general results, though not embedded in a causal framework, can be found on concentration graphs with a latent variable.

The aim of this paper is to establish a link between the two settings and to show that adapting and extending results pertaining concentration graphs can help achieving identification of principal casual effects in studies when more than one additional outcome is available. Model selection criteria are also suggested. An empirical illustrative example is provided, using data from a real social job training experiment.

Speaker: **Joe Whittaker**, University of Lancaster
Title: **Triangular symmetric elements of the entropy in graphical modelling**
Abstract:
A graphical model is defined by its joint distribution. We define the symmetric elements of the entropy for an arbitrary distribution and concentrate on the third order, the triangular elements. These are found to have interesting properties for the interpretation of the graph and leads to modest efficiency gains in graphical model search. Examples are given appropriate to a joint Gaussian distribution.