SAMSI Workshop on Large Graphical Models and Random Matrices
Abstracts

Thursday, November 9, 2006

morning

Partial inversion and partial closure of paths on graphs: two matrix operators to study properties of large systems generated over graphs

Nanny Wermuth, Chalmers/Gothenburg University/Sweden, <wermuth@math.chalmers.se>

Graphical Markov models have some advantages over other multivariate statistical models. They permit to model stepwise data generating processes with and without interventions and to work out consequences of a large model for subsets of variables. Such implications can then be compared with available background knowledge or may be used to judge seemingly inconsistent results in similar studies. Two recently developed matrix operators, one for real-valued matrices and one for binary matrices are useful for understanding and deriving properties and implications of graphical Markov models.

Separation in directed acyclic graphs: an approach based on matrix operators.

Giovanni Marchetti, University of Florence, <giovanni.marchetti@ds.unifi.it>

For directed acyclic graph models there are two known criteria to check if a conditional independence between variables A and B given C holds for distributions factorizing according to the given graph. They are called separation criteria because independence holds when the conditioning set C is a separating set, in the sense of graph theory. The first criterion introduces a new concept of d-separation but is applied to the original directed acyclic graph, while the second uses the basic notion of separation in a new induced undirected graph, called the moral graph.

We discuss an alternative approach based on several types of graphical models induced by the directe acyclic graph model, each with an associated edge matrix, i.e. a binary matrix with zeros indicating the structural independencies implied. Independencies implied by the model are assessed by checking that a particular part of the edge matrix of the induced graphical model contains all zeros. This approach suggests also a new separation criterion that is equivalent to the previous ones.

Multilevel graphical model

Anna Gottard, University of Florence, <gottard@ds.unifi.it>

Clustered data are ubiquitous in many fields of applied statistics. Multilevel modeling is one of the approaches to deal with clustered data structures to be preferred when the hierarchical structure is of interest.

In this talk, after a brief introduction to clustered data issue, I will present in detail two--level models in case of multivariate Normal distribution, highlighting a particular set of assumptions on the conditional independence structure which researchers should be aware of.

Afterwards, I'm going to introduce a class of chain graph suitable for multilevel clustered data structures. Such class of graphs are able to represent all the assumption on the conditional independence structure of a multilevel model, by using different types of node. Potentialities and limits of graphical multilevel models will be discussed.

Afternoon

Dynamic path analysis: a new approach to analyzing time-dependent covariates

Egil Ferkingstad, University of Oslo, <egil.ferkingstad@medisin.uio.no>

We introduce a general approach to dynamic path analysis. This is an extension of classical path analysis to the situation where variables may be time-dependent and where the outcome of main interest is a stochastic process. In particular we will focus on the survival and event history analysis setting where the main outcome is a counting process. Our approach will be especially fruitful for analyzing event history data with internal time-dependent covariates, where an ordinary regression analysis may fail. The approach enables us to describe how the effect of a fixed covariate partly is working directly and partly indirectly through internal time-dependent covariates. For the sequence of times of event, we define a sequence of path analysis models. At each time of an event, ordinary linear regression is used to estimate the relation between the covariates, while the additive hazard model is used for the regression of the counting process on the covariates. The methodology is illustrated using data from a randomized trial on survival for patients with liver cirrhosis.

On graphical modelling of data affected by sample selection

Elena Stanghellini, University of Perugia, <stanghel@stat.unipg.it>

The talk will focus on the use of conditional independence models with data affected by sample selection. We first review some results on distortion in linear regression coefficients induced by truncation, which is a particular from of selection. We present examples where the a priori knowledge of some conditional independencies allows to disentangle the distortions induced by truncation from the ones induced by latent variables therefore permitting to simplify the estimation of the parameters of interest. Connections to the existing models of sample selection are also established and examples are given to show how the above results on truncation can be of use in modelling data affected by other forms of sample selection, such as censoring

Friday, November 10, 2006

Morning

Derived variables

David Cox, Oxford University, <david.cox@nuffield.oxford.ac.uk>

Often the variables used for analysis and interpretation are combinations of more basic observations, sometimes called pointer readings. A familiar example is the body mass index, weight divided by the square of height. Some implications of graphical Markov models for the construction and analysis of derived variables are outlined.

A local, first-order characterization of omitted variable bias for propensity-stratified data

Ben Hansen, University of Michigan, <bbh@umich.edu>

One statistical adjustment used in observational studies is to match or stratify on an estimate of the propensity score, the conditional probability of assignment to treatment given available covariates. If potential responses and treatment assignments are conditionally independent given these covariates, then theory and practical experience suggest that the adjustment removes much of the bias due to nonrandom assignment, so that one is left with a study that, for statistical purposes, resembles an experiment randomized within the strata.

However, one must be alert to the possibility that one or more covariates omitted from the propensity adjustment would be needed to establish the necessary conditional independence. This talk outlines an asymptotic analysis of matching or stratification on estimated propensities, with some implications for omitted variable bias that is "local", in a sense to be explained. The framework appears to have relevance both to adjustment strategies based on propensity score stratification alone and to adjustment strategies combining it with other techniques, such as (perhaps) graphical models.

The framework of principal stratification for partially controlled studies

Constantine Frangakis, John Hopkins University, <cfrangak@jhsph.edu>

Our ability to examine treatment effects is often restricted to partially controlled studies, i.e., studies that control only part of the mechanism that assigns the treatments. Such studies have been typically analyzed with the standard framework of instrumental variables. However, the setting of many partially controlled studies suggests that the validity of the assumptions, and even the goals of standard instrumental variables are questionable. In this presentation, we use as examples three problems to review the role that a richer framework - principal stratification - has for (1) employing more realistic assumptions; (2) setting better goals; and (3) indicating better designs: (1) Experiments to estimate the effect of School Choice, faced with combined noncompliance to treatment assignment and incomplete follow-up data; (2) Experiments to estimate the effect of a controlled treatment on an outcome for units for which the controlled treatment does not affect an intermediate, uncontrolled factor, with application from the literature on recent vaccines trials; (3) Studies to estimate the relation that a health characteristic before a critical event (e.g., injury) has to the risk of death from the critical event, but where the exposure can be missing precisely for people who die just after the critical event.

Identification of minimal sets of covariates for matching estimators

Xavier de Luna, University of Umea, <xavier.deluna@stat.umu.se>

The Neyman-Rubin model (potential outcome framework) and the associated matching estimators have become increasingly popular, because they allow for the non-parametric estimation of average treatment effects. Like parametric models (e.g., ANCOVA), matching estimators control for a set of covariates (pre-treatment characteristics) in order to estimate the effect of a non-randomized treatment. However, unlike regression models, the selection of the covariates to be used with matching estimators has attracted little attention in the literature. This talk discusses why, when using matching estimators, the set of covariates used has to be "minimal". A set of covariates is said minimal if it cannot be reduced without violating the assumptions of the Neyman-Rubin model. Moreover, sufficient conditions are given for the identifiability of a minimal set of covariates. In order to obtain such conditions we use graphical models to impose restrictions on the set of conditional independence statements holding for the random variables involved. Finally, data-driven methods for the selection of the covariates are discussed.

afternoon

Representing equivalence classes of DAG models in the presence of selection and latent variables

Ayesha Ali, University of Guelph, <aali@uoguelph.ca>

Maximal ancestral graphs provide a way of encoding the independence relations that arise among the observed variables of DAG models with selection and latent variables. Here, we present a join operation on maximal ancestral graphs that gives rise to a unique representation for Markov equivalent maximal ancestral graphs. This equivalence class representative is analogous to the essential graph for DAGs. We also present a set of orientation rules that construct the equivalence class representative given a single member (ancestral graph) of the equivalence class. These results may be useful in model selection when one is interested in searching across equivalence classes of ancestral graphs.

Partial mapping, partial inversion and joint-response chain graphs

Michael Wiedenbeck, Mannheim Center of Social Surveys, <wiedenbeck@zuma-mannheim.de>

When a mathematical relation is inverted the picture is exchanged with the argument. The concept of partial mapping and partial inversion is introduced and discussed. Then, partial inversion is applied to nonsingular covariance matrices to derive linear least squares regression coefficients, conditional covariances and inverse marginal covariance matrices. Another application is the block-triangular decomposition of covariance matrices which is related to different types of joint-response chain graphs.

A new technic to prove Markov equivalence.

Oscar Hammar, Chalmers/Gothenburg University/Sweden, <oscham@math.chalmers.se>

Markov equivalence between certain types of chain graphs have been proven in the literature using path criteria. We use a matrix formulation for induced edges to simplify proofs a directed acyclic graph to chain graphs and to tackle new problems.

Saturday, November 11, 2006

morning

High-Dimensional Networks as Conduits for Uncertain Information

Arthur Dempster, Harvard University, <dempster@fas.harvard.edu>

After discussing what network models are and are not, and sketching several applied areas of contemporary interest and importance, I argue that DS (ÒDempster-ShaferÓ) methodology is a natural tool for modeling and analysis of complex networks because (1) information needs to be propagated and fused from widely dispersed regions of the network; and (2) information always remains completely absent regarding much of the network.

DS theory provides a natural and flexible extension of Bayesian methods, impartially mixing probabilistic uncertainty and logical degrees of ÒdonÕt knowÓ. Important elements of the theory were published, 40, 30, and 20 years ago, but computational feasibility is just now coming on line.

The conjugate prior for discrete hierarchical models

Helen Massam, Unviversity of York, <massamh@mathstat.yorku.ca>

Abstract: In the Bayesian analysis of contingency table data, the selection of a prior distribution for either the log-linear parameters or the cell probabilities parameter is a major challenge. Though the conjugate prior on cell probabilities has been defined by Dawid and Lauritzen (1993) for decomposable graphical models, it has not been identified for the larger class of graphical models Markov with respect to an arbitrary undirected graph or for the even wider class of hierarchical log-linear models. Working with the log-linear parameters used by GLIM, we first define the conjugate prior for these parameters and then derive the induced prior for the cell probabilities: this is done for the general class of hierarchical log-linear models. We show that the conjugate prior has all the properties that one expects from a prior: notational simplicity, ability to reflect either no prior knowledge or a priori expert knowledge, a moderate number of hyperparameters and mathematical convenience. It also has the strong hyper Markov property which allows for local updates within prime components for graphical models.

Robust model selection for graphical models

Sonja Kuhnt, University og Eindhoven, <S.Kuhnt@TUe.nl>

In every data set observations can occur which deviate from the pattern presented by the main part of the data. Such so-called outliers can severely influence the result of a model selection process. We will provide a formal definition of outliers with respect to CG-distributions, capturing the perception of outliers as surprising observations. Based on the usual likelihood equations we define a general class of modified likelihood estimators which allow for robust parameter estimation. Finally the combination of one-step outlier identification with well-known model selection strategies is discussed with a focus on undirected graphical models for mixed variables.

Active Structure Search using Interventions

Fredrick Eberhard, Carnegie Mellon University, <fde@cmu.edu>

The causal Bayes net framework (Pearl, 2000; Spirtes et al, 2000) provides a representation of causal structure that captures both the probabilistic relations between variables and the effect of interventions on variables. I will present a general model that captures a broad range of different types of interventions, which imply different experimental search strategies. I will provide worst case bounds on the number of experiments necessary and sufficient to recover the true causal structure and describe some work in progress on identifying the shortest sequence of experiments.