SLOW MONDAY in ALGEBRAIC STATISTICS
September 18, 2017
Starting time 10am

Department of Mathematics, University of Genova
Room 705

ORGANIZERS

Fabio Rapallo, Università del Piemonte Orientale
Eva Riccomagno, Università di Genova

PROGRAMME

10.00 – 10.10 Eva Riccomagno: Welcome and introduction
10.10 – 10.50 Gherardo Varando: The algebra of generative classifiers
10.50 – 11.30 Christiane Görgen: Graphical statistical models which are not toric varieties
11.30 – 12.10 Luigi Malagò: Variational autoencoders, graphical models

12.10 – 14.00 Lunch

14.00 – 14.40 Henry Wynn: Optimal experimental design that minimizes the width of simultaneous confidence bands
14.40 – 15.20 Jim Smith: Finding promising causal hypotheses without graphs
15.20 – 16.00 Giovanni Pistone: Modal logic and Algebraic Statistics

PARTICIPANTS

Christiane Görgen, Max Planck Institute for Mathematics Leipzig
Luigi Malagò, Romanian Institute of Science and Technology
Jim Q. Smith, The University of Warwick
Gherardo Varando, Universidad Politecnica de Madrid
Henry P. Wynn, London School of Economics and Political Sciences

Roberto Fontana, Politecnico di Torino
Giovanni Pistone, Collegio Carlo Alberto, Moncalieri, Italy
Fabio Rapallo, Università del Piemonte Orientale

Claudia Fassino, Università di Genova
Elena Pesce, Università di Genova
Eva Riccomagno, Università di Genova
Maria-Piera Rogantin, Università di Genova
Giacomo Siri, Università di Genova

ABSTRACTS

CHRISTIANE GÖRGEN**: Graphical statistical models which are not toric varieties**

Abstract. One of the fundamental insights in Algebraic Statistics is that exponential families 'are' toric varieties. A full characterisation of the algebraic geometry of a special class of these, called decomposable graphical models, was given by Geiger et al. (2006). We now investigate a more recent statistical model represented by a coloured probability tree and called a staged tree (Collazo et al., 2017). Staged tree models include decomposable graphical models as a special case but in their algebro-geometric description sum-to-1 conditions on the probability simplex cannot be ignored. These hence constitute a wider class of models - and of varieties - which are toric only under very special conditions.
This is ongoing work with Eliana Duarte from the Max-Planck-Institute of Mathematics in the Sciences and with Manuele Leonelli from the University of Glasgow. I will show that staged tree models are curved exponential models and will give a full description of these models in terms of semi-algebraic sets.

LUIGI MALAGÒ: **Variational Autoencoders - Current Research and Future Directions**

Abstract. In this presentation we give an overview of variational inference in deep learning, by reviewing the general algorithms for training Variational Autoencoders, their limitations and the challenges they pose, which lead to different research directions that can be found in the vast literature. We continue with a presentation of our previous and current research, such as the use of Gaussian Graphical Models and the analysis of the geometry of the latent space. The presentation concludes with research questions and directions for the future, related to the study of the geometry of the latent space and the use of different bounds derived from different divergences.

GIOVANNI PISTONE: **Modal logic and Algebraic Statistics**

Abstract. In the present paper we consider modal propositional logic with a finite number of propositional variables and look for those constraints which are imposed to the propositions of a special type by the structure of the relevant Kripke's frame. We translate the

usual language of modal propositional logic in terms of notions of commutative algebra, namely polynomial rings, ideal and bases of ideals. We use extensively the perspective obtained in previous papers in Algebraic Statistics.

JIM SMITH: **Finding promising causal hypotheses without graphs**

Abstract. Multivariate models can often be described using a Bayesian Network. When this is so, various authors have asserted that certain topologies of these directed graphs might indicate causal relationship between its variables. But can this type of technology be extended so that it can be used to find promising causal relationships when the variables that will become part of any causal relationship are yet to be defined? In this talk I will argue that this might indeed be the case. After discussing some basic causal discovery algorithms for BNs I will suggest analogues that apply to the much wider class of chain event graph.

GHERARDO VARANDO: **The algebra of generative classifiers**

Abstract. In this ongoing work we study algebraic and geometric aspects of generative binary classifiers over discrete predictor variables. We prove that every conditional independence statement holding in the underlying probability is equivalent to a set of linear equivalences for the induced discrimination function. Moreover we prove that the former is equivalent to the odds ratios of the conditional probabilities given the class values being identical. We thus study the spaces of constant odds ratios and we present some examples for

HENRY WYNN: **Optimal experimental design that minimizes the width of simultaneous confidence bands**

Abstract. We propose an optimal experimental design for a curvilinear regression model that minimizes the bandwidth of simultaneous confidence bands. Simultaneous confidence bands for nonlinear regression are constructed by evaluating the volume of a tube about a curve that is defined as a trajectory of a regression basis vector (Naiman, 1986). The proposed criterion is constructed based on the volume of a tube, and the corresponding optimal design is referred to as the minimum-volume optimal design. For Fourier and weighted polynomial regressions, the problem is formalized as one of minimization over the cone of Hankel positive definite matrices, and the criterion to minimize is expressed as an elliptic integral. We show that the Mobius group keeps our problem invariant, and hence, minimization can be conducted over cross-sections of orbits. We demonstrate that for the weighted polynomial regression and the Fourier regression with three bases, the minimum-volume optimal design forms an orbit of the Mobius group containing D-optimal designs as representative elements.