



Host: Univ.-Prof. Christoph Kuzmics, PhD

INVITATION

Prof. Dr. Maria-Pia Victoria Feser
(University of Bologna)

On the topic:

"Fast M-Estimation of Generalized Linear Latent Variable Models in High Dimensions"

Dimension reduction for high dimensional data is an important and challenging task, relevant to both machine learning and statistical applications. Generalized Linear Latent Variable Models (GLLVMs) provide a probabilistic alternative to matrix factorization when the data are of mixed types, whether discrete, continuous, or a mixture of both. They achieve the reduction of dimensionality by mapping the correlated multivariate data to so-called latent variables, defined in a lower-dimensional space. The benefit of GLLVMs is twofold: the latent variables can be estimated and used as features to be embedded in another model, and the model parameters themselves are interpretable and provide meaningful indications on the very structure of the data. Moreover, GLLVM can naturally be extended to dynamic processes such as those used to model longitudinal data. However, with a likelihood based approach, GLLVM's estimation represents a tremendous challenge for even moderately large dimensions, essentially due to the multiple integrals involved in the likelihood function. Numerous methods based on approximations of this latter have been proposed: Laplace approximation, adaptive quadrature, or, recently, extended variational approximation. For GLLVMs, however, these methods do not scale well to high dimensions, and they may also introduce a large bias in the estimates. In this paper, we consider an alternative route, which consists in proposing an alternative estimator, based on drastically simplified estimating equations, complemented with a numerically efficient bias reduction methods in order to recover a consistent estimator for the GLLVM parameters. The resulting estimator is an M-estimator, which has a negligible efficiency loss compared to the (exact) MLE. For larger data sets, the proposed M-estimator, whose computational burden is linear in npq , remains applicable when the state-of-the-art method fails to converge. To compute the M-estimator, we propose to use a stochastic approximation algorithm.

Date: October 24, 2023

Starting time: 4.30 p.m.

Location: HS 111.21/Beethovenstraße 8/2

All professors, lecturers, assistants and students are cordially invited to this lecture!