Paloma Uribe (USP)

Title: Sparse and time-varying covariance modeling

Abstract: In many areas such as neuroscience, energy planning and finance there has been a growing interest in developing computationally fast methods that can estimate dependency of high-dimensional multivariate time series data. Nevertheless, the estimation of a covariance matrix based on high-dimension data is still an open problem. Another important issue is whether these covariances exhibit time varying patterns. In this paper, we discuss several bayesian regularization methods based on shrinkage and selection priors and estimate sparse covariance matrices using the modified Cholesky decomposition. Our first model considers dynamics only for the variances (stochastic volatility) and uses the Normal-Gamma prior for shrinking the regression coefficients that compose the Cholesky factor. The second model considers homoscedastic errors and time varying regression coefficients generated by a dynamic version of the Normal Mixture of Inverse Gammas (NMIG) hierarchical prior, which accommodates time varying sparsity.

Guilherme Ludwig (UNICAMP)

Title: Static and Roving Sensor Data Fusion for Hazard Mapping

Abstract: Rapid technological advances have drastically improved the data collection capacity in occupational exposure assessment. However, advanced statistical methods for analyzing such data and drawing proper inference remain limited. The objectives of this paper are (1) to provide new spatio-temporal methodology that combines data from both roving and static sensors for data processing and hazard mapping across space and over time in an indoor environment, and (2) to compare the new method with the current industry practice, demonstrating the distinct advantages of the new method and the impact on occupational hazard assessment and future policy making in environmental health as well as occupational health. A novel spatio-temporal model with continuous index in both space and time is proposed, and a profile likelihood based model fitting procedure is developed that allows fusion of the two types of data. To account for potential differences between the static and roving sensors, we extend the model to have non-homogenous measurement error variances. Our methodology is applied to a case study conducted in an engine test facility and dynamic hazard maps are drawn to show features in the data that would have been missed by existing approaches, but are captured by the new method.

Rafael Izbicki (UFSCAR)

Title: Converting High-Dimensional Regression to High-Dimensional Nonparametric Conditional Density Estimation

Abstract: There is a growing demand for nonparametric conditional density estimators (CDEs) in fields such as astronomy and economics. In astronomy, for example, one can dramatically improve estimates of the parameters that dictate the evolution of the Universe by working with full conditional densities instead of regression (i.e., conditional mean) estimates. More generally, standard regression falls short in any prediction problem where the distribution of the response is more complex with multi-modality, asymmetry or heteroscedastic noise. Nevertheless, much of the work on high-dimensional inference concerns regression and classification only, whereas research on density estimation has lagged behind. Here we propose FlexCode, a fully nonparametric approach to conditional density estimation that reformulates CDE as a non-parametric orthogonal series problem where the expansion coefficients are estimated by regression. By taking such an approach, one can efficiently estimate conditional densities and not just expectations in high dimensions by drawing upon the success in high-dimensional regression. Depending on the choice of regression procedure, our method can adapt to a variety of challenging high-dimensional settings with different structures in the data (e.g., a large number of irrelevant components and nonlinear manifold structure) as well as different data types (e.g., functional data, mixed data types and sample sets). We study the theoretical and empirical performance of our proposed method, and we compare our approach with traditional conditional density estimators on simulated as well as real-world data, such as photometric galaxy data, Twitter data, and line-of-sight velocities in a galaxy cluster.

Flávio Ziegelmann (UFRGS)

Title: Dynamics of Financial Returns Densities: A Functional Approach Applied to the Bovespa Intraday Index

Abstract: We model the stochastic evolution of probability density functions (pdf's) of IBOVESPA intraday returns over business days, in a functional time series framework. We find evidence that the pdf's dynamic structure reduces to a vector process lying on a two-dimensional space. Our main contributions are as follows. First, we provide further

insight on the finite-dimensional decomposition of the curve process: it is shown that its evolution can be interpreted as a dynamic dispersion-symmetry shift. Second, we provide an application to realized volatility forecasting, with a forecasting ability comparable to HAR Realized Volatility models in the Model Confidence Set framework.

Eufrásio de Andrade Lima Neto (UFPB)

Title: A robust regression method based on exponential-type kernel functions

Abstract: The use of robust regression methods is common in practical situations due to the presence of outliers. This work proposes a robust regression method that re-weighted the outliers observations considering type-exponential kernel functions. The convergence of the parameter estimate algorithm is guaranteed with a low computational cost. A comparative study between the proposed regression method (ETKRR) against some classical robust approaches and the OLS method is considered. We have considered synthetic datasets with X-axis outliers, Y-axis outliers and leverage points, in a Monte Carlo simulation framework with different sample sizes and percentage of outliers. The results have demonstrated that the ETKRR approach presented a competitive (or best) performance in simulation scenarios that are similar to those found in real problems. Applications to real datasets has showed the usefulness of the proposed method.