Abstracts

Contents

Updated: August 20, 2024

Tony Cai
University of Pennsylvania, USA

Transfer Learning for Funchtional Data Analysis

TBA

Back to Contents Page →

Jinyuan Chang
Southwestern University of Finance and Economics, China

On the modelling and prediction of high-dimensional functional time series

We propose a two-step procedure to model and predict high-dimensional functional time series, where the number of function-valued time series p is large in relation to the length of time series n. Our first step performs an eigenanalysis of a positive definite matrix, which leads to a one-to-one linear transformation for the original high dimensional functional time series, and the transformed curve series can be segmented into several groups such that any two subseries from any two different groups are uncorrelated both contemporaneously and serially. Consequently in our second step those groups are handled separately without the information loss on the overall linear dynamic structure. The second step is devoted to establishing a finite-dimensional dynamical structure for all the transformed functional time series within each group.

Furthermore the finite-dimensional structure is represented by that of a vector time series. Modelling and forecasting for the original high-dimensional functional time series are realized via those for the vector time series in all the groups. We investigate the theoretical properties of our proposed methods, and illustrate the finite-sample performance through both extensive simulation and three real datasets.

Back to Contents Page →

Maria Grith
Erasmus University Rotterdam, Netherlands

Group factors in single stock options

We consider a linear functional panel model in which the response variables are functions, and the regressors are scalar common factors. The functional response introduces a third dimension in the panel alongside the cross-section and time dimensions. We assume that the coefficient functions are characterized by a multilevel latent group pattern of heterogeneity. For discretely observed functions, we use penalized-sieve estimation. We estimate the group memberships and functional coefficients using an estimator that minimizes a least squares criterion with respect to all possible groupings of the crosssectional response functions when the number of groups is known. We provide conditions under which our estimators are consistent as the observations in the three dimensions of the panel tend to infinity, and we develop inference methods. Additionally, we show that using a functional setup improves clustering accuracy and convergence rates of the group-specific functional coefficients. The gains are enhanced if membership is factor-invariant. We also propose a generalized information criterion for estimating the number of groups when it is unknown. We evaluate the performance of our method in a simulation with both densely and sparsely observed functional responses. Finally, we apply our approach to identify grouped patterns of unobserved heterogeneity in a panel of single-stock options implied volatility surfaces.

Back to Contents Page →

Wanjie Wang
National University of Singapore, Singapore

Recovery of Time Labels for Noisy Dynamic Data

The dynamic system is frequently observed in science, with plenty of works in many directions. We observe $X_i \in R^p$ at time points t_i, $1 <= i <= N$. Based on X_i, we want to reconstruct the system $X(t)$. However, in some applications, the time labels t_i's are not available. For example, in cyro-EM data, many 2D images of one molecule that is under molecule dynamic are captured. To reconstruct the 3D structure and understand the dynamic, we should know the ``stage" of each image, i.e. t_i. When p is large, the noise is large and the problem becomes very challenging. In this work, we present a spectral method to recover t_i. With manifold learning, we present the consistency of our method, even in the presence of high-dimensional noise. It is the first work to guarantee the consistency of the spectral seriation algorithm for complicated cases under the high-dimensional noise. In numerical analysis, we show its power.

Back to Contents Page →

Jin-Ting Zhang
National University of Singapore, Singapore

Two-Sample Tests for Equal Distributions in Separable Metric Spaces: A Unified Semimetric-Based Approach

With the advancement of data collection techniques, researchers frequently encounter complex data objects within separable metric spaces across various domains. One common interest lies in determining whether two groups of complex data objects originate from the same population. This paper introduces and examines a fast and accurate unified semimetric-based approach designed to tackle this challenge. The approach exhibits broad applicability across a wide range of research areas, such as bioinformatics, audiology, environmentology, finance, and more. It effectively identifies differences between the distributions of two complex datasets, including both high-dimensional data and functional data. The asymptotic null and alternative distributions of the proposed test statistic are established. Unlike the permutation approach, a unified, rapid and precise method to approximate the null distribution is described. Furthermore, the proposed test is shown to be root-n consistent. Numerical results are presented for illustrating the excellent performance of the proposed test in terms of size control, power, and computational cost. Additionally, the applications of the proposed test are showcased through examples involving both high-dimensional data and functional data.

Back to Contents Page →