A list of forthcoming papers can be found online at https://www.bernoullisociety.org/index.php/publications/bernoulli-journal/bernoulli-journal-papers
CONTENTS

(continued)

GUO, S. and QIAO, X.  
On consistency and sparsity for high-dimensional functional time series with application to autoregressions  
451

VAŠKEVIČIUS, T. and ZHIVOTOVSKIY, N.  
Suboptimality of constrained least squares and improvements via non-linear predictors  
473

CONT, R. and DAS, P.  
Quadratic variation and quadratic roughness  
496

LEE, C.Y. and XIAO, Y.  
Chung-type law of the iterated logarithm and exact moduli of continuity for a class of anisotropic Gaussian random fields  
523

HEINE, K. and BURROWS, D.  
Multilevel bootstrap particle filter  
551

WANG, T.  
Exponential and strong ergodicity for one-dimensional time-changed symmetric stable processes  
580

MALLEIN, B. and SHI, Q.  
A necessary and sufficient condition for the convergence of the derivative martingale in a branching Lévy process  
597

GASS, L.  
Almost-sure asymptotics for Riemannian random waves  
625

EFTEKHARI, H., BANERJEE, M. and RITOV, Y.  
Design of c-optimal experiments for high-dimensional linear models  
652

DEBALY, Z.-M. and TRUQUET, L.  
Multivariate time series models for mixed data  
669

TORRISI, G.L., GARETTO, M. and LEONARDI, E.  
Bootstrap percolation on the stochastic block model  
696

TSCHIDERER, B. and CHUN YEUNG, L.  
A trajectorial approach to relative entropy dissipation of McKean–Vlasov diffusions: Gradient flows and HWBI inequalities  
725

RÖCKNER, M. and ZHAO, G.  
SDEs with critical time dependent drifts: Weak solutions  
757

DING, Y., PENG, Q. and XIAO, Y.  
Linear multifractional stable sheets in the broad sense: Existence and joint continuity of local times  
785

NG, T.L.J. and ZAMMIT-MANGION, A.  
Non-homogeneous Poisson process intensity modeling and estimation using measure transport  
815

PUCETTI, G. and RÜSCHENDORF, L.  
General construction and classes of explicit $L^1$-optimal couplings  
839
BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Aims and Scope
Issued four times per year, BERNOULLI is the flagship journal of the Bernoulli Society for Mathematical Statistics and Probability. The journal aims at publishing original research contributions of the highest quality in all subfields of Mathematical Statistics and Probability. The main emphasis of Bernoulli is on theoretical work, yet discussion of interesting applications in relation to the proposed methodology is also welcome.

Bernoulli Society for Mathematical Statistics and Probability
The Bernoulli Society was founded in 1973. It is an autonomous Association of the International Statistical Institute, ISI. According to its statutes, the object of the Bernoulli Society is the advancement, through international contacts, of the sciences of probability (including the theory of stochastic processes) and mathematical statistics and of their applications to all those aspects of human endeavour which are directed towards the increase of natural knowledge and the welfare of mankind.

Meetings: https://www.bernoullisociety.org/meetings
The Society holds a World Congress every four years; more frequent meetings, coordinated by the Society’s standing committees and often organised in collaboration with other organisations, are the European Meeting of Statisticians, the Conference on Stochastic Processes and their Applications, the CLAPEM meeting (Latin-American Congress on Probability and Mathematical Statistics), the European Young Statisticians Meeting, and various meetings on special topics – in the physical sciences in particular. The Society, as an association of the ISI, also collaborates with other ISI associations in the organization of the biennial ISI World Statistics Congresses (formerly ISI Sessions).

Executive Committee
Detailed information about the members of the Executive Committee can be found on https://www.bernoullisociety.org/who-is-who

The papers published in Bernoulli are indexed or abstracted in Mathematical Reviews (MathSciNet), Zentralblatt MATH (zbMATH Open), Science Citation Index Expanded (Web of Science), SCOPUS and Google Scholar.

©2023 International Statistical Institute/Bernoulli Society

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the Publisher.

In 2023 Bernoulli consists of 4 issues published in February, May, August and November.
Conformal prediction: A unified review of theory and new challenges

MATTEO FONTANA\textsuperscript{1,2,3,c}, GIANLUCA ZENI\textsuperscript{1,a} and SIMONE VANTINI\textsuperscript{1,b}

\textsuperscript{1}MOX-Department of Mathematics, Politecnico di Milano, Italy. \textsuperscript{a}gianluca.zeni@mail.polimi.it, \textsuperscript{b}simone.vantini@polimi.it
\textsuperscript{2}Department of Management, Economics and Industrial Engineering, Politecnico di Milano, Italy
\textsuperscript{3}now at European Commission, Joint Research Centre (JRC), Ispra, Italy. \textsuperscript{c}matteo.fontana@ec.europa.eu

In this work we provide a review of basic ideas and novel developments about Conformal Prediction — an innovative distribution-free, non-parametric forecasting method, based on minimal assumptions — that is able to yield in a very straightforward way prediction sets that are valid in a statistical sense also in the finite sample case. The discussion provided in the paper covers the theoretical underpinnings of Conformal Prediction, and then proceeds to list the more advanced developments and adaptations of the original idea.

Keywords: Conformal prediction; nonparametric statistics; prediction intervals; review

References


In this paper, we explore the two star Exponential Random Graph Model, which is a two parameter exponential family on the space of simple labeled graphs. We introduce auxiliary variables to express the two star model as a mixture of the $\beta$ model on networks. Using this representation, we study asymptotic distribution of the number of edges, and the sampling variance of the degrees. In particular, the limiting distribution for the number of edges has similar phase transition behavior to that of the magnetization in the Curie-Weiss Ising model of Statistical Physics. Using this, we show existence of consistent estimates for both parameters. Finally, we prove that the centered partial sum of degrees converges as a process to a Brownian bridge in all parameter domains, irrespective of the phase transition.

Keywords: ERGM; auxiliary variables; phase transition; consistent estimation; two-star

References


Template matching with ranks

ERY ARIAS-CASTRO\textsuperscript{a} and LIN ZHENG\textsuperscript{b}

Department of Mathematics, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0112, USA. \textsuperscript{a}eariascastro@ucsd.edu, \textsuperscript{b}liz176@ucsd.edu

We consider the problem of matching a template to a noisy signal. Motivated by some recent proposals in the signal processing literature, we suggest a rank-based method and study its asymptotic properties using some well-established techniques in empirical process theory combined with Hájek’s projection method. The resulting estimator of the shift is shown to achieve a parametric rate of convergence and to be asymptotically normal. Some numerical simulations corroborate these findings.

Keywords: Matched filter; template matching; scan statistics; Spearman rank correlation; empirical processes; minimax optimality

References


Mixing time guarantees for unadjusted Hamiltonian Monte Carlo

NAWAF BOU-RABEE$^{1,a}$ and ANDREAS EBERLE$^{2,b}$

$^1$Department of Mathematical Sciences, Rutgers University Camden, 311 N 5th Street, Camden, NJ 08102 USA.

$^a$nawaf.bourabee@rutgers.edu

$^2$Institut für Angewandte Mathematik, Universität Bonn, Endenicher Allee 60, 53115 Bonn, Germany.

b eberle@uni-bonn.de

We provide quantitative upper bounds on the total variation mixing time of the Markov chain corresponding to the unadjusted Hamiltonian Monte Carlo (uHMC) algorithm. For two general classes of models and fixed time discretization step size $h$, the mixing time is shown to depend only logarithmically on the dimension. Moreover, we provide quantitative upper bounds on the total variation distance between the invariant measure of the uHMC chain and the true target measure. As a consequence, we show that an $\epsilon$-accurate approximation of the target distribution $\mu$ in total variation distance can be achieved by uHMC: (i) for a broad class of models with $O\left(\frac{d^3}{4} \epsilon^{-1/2} \log(d/\epsilon)\right)$ gradient evaluations; and (ii) for mean field models with weak interactions with $O\left(\frac{d^{1/2}}{\epsilon^{-1/2}} \log(d/\epsilon)\right)$ gradient evaluations. The proofs are based on the construction of successful couplings for uHMC that realize the upper bounds.

Keywords: MCMC; Hamiltonian Monte Carlo; mixing time; couplings; variational integrators

References


Mixing time guarantees for uHMC


High dimensional generalized linear models for temporal dependent data

YUEFENG HAN\textsuperscript{1,a}, RUEY S. TSAY\textsuperscript{2,b} and WEI BIAO WU\textsuperscript{3,c}

\textsuperscript{1}Department of Statistics, Rutgers University, Piscataway, NJ 08854, USA. \textsuperscript{a}yuefeng.han@rutgers.edu
\textsuperscript{2}Booth School of Business, University of Chicago, Chicago, IL 60637, USA. \textsuperscript{b}ruey.tsay@chicagobooth.edu
\textsuperscript{3}Department of Statistics, University of Chicago, Chicago, IL 60637, USA. \textsuperscript{c}wbwu@galton.uchicago.edu

High dimensional generalized linear models are widely applicable in many scientific fields with data having heavy tails. However, little is known about statistical guarantees on the estimates of such models in a time series setting. In this article, we establish statistical error bounds and support recovery guarantees of the classical $\ell_1$ regularized procedure for generalized linear model with temporal dependent data. We also propose a new robust $M$-estimator for high dimensional time series. Properties of the proposed robust procedure are investigated both theoretically and numerically. As an extension, we introduce a robust estimator for linear regression and show that the proposed robust estimator achieves nearly the optimal rate as that for i.i.d sub-Gaussian data. Simulation results show that the proposed method performs well numerically in the presence of heavy-tailed and serially dependent covariates and/or errors, and it significantly outperforms the classical Lasso method. For applications, we demonstrate, in the supplementary material, the regularized robust procedure via analyzing high-frequency trading data in finance.

Keywords: High dimensional analysis; time series analysis; generalized linear model; robust estimation; support recovery

References


We define and prove limit results for a class of dominant Pólya sequences, which are randomly reinforced urn processes with color-specific random weights and unbounded number of possible colors. Under fairly mild assumptions on the expected reinforcement, we show that the predictive and the empirical distributions converge almost surely (a.s.) in total variation to the same random probability measure \( \tilde{P} \); moreover, \( \tilde{P}(D) = 1 \) a.s., where \( D \) denotes the set of dominant colors for which the expected reinforcement is maximum. In the general case, the predictive probabilities and the empirical frequencies of any \( \delta \)-neighborhood of \( D \) converge a.s. to one. That is, although non-dominant colors continue to be regularly observed, their distance to \( D \) converges in probability to zero. We refine the above results with rates of convergence. We further hint potential applications of dominant Pólya sequences in randomized clinical trials and species sampling, and use our central limit results for Bayesian inference.

Keywords: Reinforced processes; Pólya sequences; random probability measures; species sampling; Bayesian nonparametrics

References


Dominant Pólya sequences


On admissible estimation of a mean vector when the scale is unknown

YUZO MARUYAMA\textsuperscript{1,a} and WILLIAM E. STRAWDERMAN\textsuperscript{2,b}

\textsuperscript{1}Graduate School of Business Administration, Kobe University, Kobe, Japan. \textsuperscript{a}maruyama@port.kobe-u.ac.jp
\textsuperscript{2}Department of Statistics and Biostatistics, Rutgers University, New Brunswick, NJ, USA. \textsuperscript{b}straw@stat.rutgers.edu

We consider admissibility of generalized Bayes estimators of the mean of a $p$-variate normal distribution when the scale is unknown, and the loss is quadratic. The priors considered put the improper invariant prior on the scale while the prior on the mean has a hierarchical normal structure conditional on the scale. This conditional hierarchical prior is indexed by a hyperparameter, $a$. In earlier studies, the authors established admissibility/in-admissibility of the generalized Bayes estimator under the proper/improper conditional prior ($a > -1 / a < -2$), respectively. In this paper we complete the admissibility/inadmissibility characterization for this class of priors by establishing admissibility for the improper conditional prior ($-2 \leq a \leq -1$). This boundary, $a = -2$, with admissibility for $a \geq -2$ and inadmissibility for $a < -2$ corresponds exactly to that in the known scale case for this class of conditional priors, and which follows from Brown’s 1971 paper. As a notable benefit of this enlargement of the class of admissible generalized Bayes estimators, we give admissible and minimax estimators for $p \geq 3$ as opposed to an earlier study which required $p \geq 5$. In one particularly interesting special case, we establish that the joint Stein prior for the unknown scale case leads to a minimax admissible estimator for $p \geq 3$.

Keywords: Admissibility; Bayes estimators; Minimaxity

References


On a projection-based class of uniformity tests on the hypersphere

EDUARDO GARCÍA-PORTUGUÉS¹,a, PAULA NAVARRO-ESTEBAN²,b and JUAN A. CUESTA-ALBERTOS²,c

¹Departamento de Estadística, Universidad Carlos III de Madrid, Leganés, Spain. ²Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria, Santander, Spain. ³edgarcia@est-econ.uc3m.es, ⁴paula.navarro@unican.es, ⁵juan.cuesta@unican.es

We introduce a projection-based class of uniformity tests on the hypersphere. The class employs statistics that integrate, along all possible directions, a weighted quadratic discrepancy between the empirical cumulative distribution of the projections and the projected uniform distribution. Simple expressions for several test statistics are obtained for the circle and the sphere, as well as relatively tractable forms for higher dimensions. Despite their different origins, variants of the proposed class are shown to contain and be contained in variants of the Sobolev class of uniformity tests. Our new class proves itself advantageous by allowing the derivation of new tests that neatly extend the circular-only tests by Watson, Ajne, and Rothman, and by introducing the first instance of an Anderson–Darling-like test for directional data. We obtain usable asymptotic distributions and the local asymptotic optimality against certain alternatives of the new tests. A simulation study evaluates the theoretical findings and provides evidence that the new testing proposals are competitive. An application to the study of the crater distribution on Rhea illustrates the usage of the new tests.

Keywords: Circular data; directional data; projection; uniformity

References


Donsker results for the empirical process indexed by functions of locally bounded variation and applications to the smoothed empirical process

ERIC BEUTNER\textsuperscript{1,a} and HENRYK ZÄHLE\textsuperscript{2,b}

\textsuperscript{1}Department of Econometrics and Data Science, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands. \textsuperscript{a}e.a.beutner@vu.nl

\textsuperscript{2}Department of Mathematics, Saarland University, Germany. \textsuperscript{b}zaehle@math.uni-sb.de

Recently, Radulovic and Wegkamp introduced a new technique to show convergence in distribution of the empirical process indexed by functions of bounded variation. This method of proof allows to directly extend convergence results known for the canonical empirical process to convergence in distribution of the empirical process indexed by functions of bounded variation. The purpose of this article is twofold. First, we extend the mentioned technique to index functions of \textit{locally} bounded variation. Second, and more importantly, we demonstrate that this technique provides a new approach to show convergence in distribution of the smoothed empirical process based on kernel density estimators. Using this approach we can prove to the best of our knowledge the first results on convergence in distribution of the smoothed empirical process of \textit{dependent data}. Our results cover both weak and strong dependence as well as index sets of functions of locally bounded variation. Moreover our results cover an MISE optimal choice of the bandwidth for the kernel density estimator which to some extent is the plug-in property in the Bickel–Ritov sense. In the case of i.i.d. data our results extend a seminal result of Giné and Nickl.

\textbf{Keywords:} Weak convergence; empirical process; function of locally bounded variation; smoothed empirical process; kernel smoothing; optimal bandwidth; plug-in property

\section*{References}


Donsker results for empirical and smoothed empirical processes


Rank-based testing for semiparametric VAR models: A measure transportation approach

MARC HALLIN\textsuperscript{1,a}, DAVIDE LA VECCHIA\textsuperscript{2,b} and HANG LIU\textsuperscript{3,c}

\textsuperscript{1}ECARES and Département de Mathématique, Université libre de Bruxelles CP 114/4, Avenue F.D. Roosevelt 50, B-1050 Bruxelles, Belgium. \textsuperscript{amhallin@ulb.ac.be}

\textsuperscript{2}Research Center for Statistics, University of Geneva, Bd du Pont d’Arve 40 - CH-1211 Geneva, Switzerland. \textsuperscript{bdavide.lavecchia@unige.ch}

\textsuperscript{3}International Institute of Finance, School of Management, University of Science and Technology of China, Hefei, China. \textsuperscript{chliu01@ustc.edu.cn}

We develop a class of tests for semiparametric vector autoregressive (VAR) models with unspecified innovation densities based on the recent measure-transportation-based concepts of multivariate center-outward ranks and signs. We show that these concepts, combined with Le Cam’s asymptotic theory of statistical experiments, yield novel testing procedures, which (a) are valid under a broad class of innovation densities (possibly non-elliptical, skewed, and/or with infinite moments), (b) are optimal (locally asymptotically maximin or most stringent) at selected ones, and (c) are robust against additive outliers. In order to show this, we establish, for a general class of center-outward rank-based serial statistics, a Hájek asymptotic representation result, of independent interest, which allows for a rank-based reconstruction of central sequences. As an illustration, we consider the problems of testing the absence of serial correlation in multiple-output and possibly non-linear regression (an extension of the classical Durbin-Watson problem) and the sequential identification of the order $p$ of a VAR($p$) model. A Monte Carlo comparative study of our tests and their routinely-applied Gaussian competitors demonstrates the benefits (in terms of size, power, and robustness) of our methodology; these benefits are particularly significant in the presence of asymmetric and leptokurtic innovation densities. A real-data application concludes the paper.

\textbf{Keywords:} Multivariate ranks; distribution-freeness; Hájek representation; local asymptotic normality; Durbin-Watson test; VAR order selection

\section*{References}


Cramér-type moderate deviation of normal approximation for unbounded exchangeable pairs

ZHOU-SONG ZHANG

Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546.

In Stein’s method, the exchangeable pair approach is commonly used to estimate the approximation errors in normal approximation. In this paper, we establish a Cramér-type moderate deviation theorem of normal approximation for unbounded exchangeable pairs. As applications, Cramér-type moderate deviation theorems for the sums of local statistics and general Curie–Weiss model are obtained.

Keywords: Stein’s method; exchangeable pair approach; Cramér-type moderate deviation; sums of local statistics; general Curie–Weiss model

References


Variational formulas for asymptotic variance of general discrete-time Markov chains

LU-JING HUANG¹,a and YONG-HUA MAO²,b

¹School of Mathematics and Statistics, Fujian Normal University, 350007 Fuzhou, P.R. China. 
²School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, 100875 Beijing, P.R. China. 
a huanglj@fjnu.edu.cn 
b maoyh@bnu.edu.cn

The asymptotic variance is an important criterion to evaluate the performance of Markov chains, especially for the central limit theorems. We give the variational formulas for the asymptotic variance of discrete-time (non-reversible) Markov chains on general state space. The variational formulas provide many applications, extending the classical Peskun’s comparison theorem to non-reversible Markov chains, and obtaining several comparison theorems between Markov chains with various perturbations.

Keywords: Markov chain; asymptotic variance; variational formula; non-reversible; Peskun’s theorem; comparison theorem

References

Variational formulas for asymptotic variance


Stratified incomplete local simplex tests for curvature of nonparametric multiple regression

YANGLEI SONG\textsuperscript{1,a}, XIAOHUI CHEN\textsuperscript{2,b} and KENGO KATO\textsuperscript{3,c}

\textsuperscript{1}Department of Mathematics and Statistics, Queen's University, Jeffery Hall, Kingston, ON, Canada, K7L 3N6.
\textsuperscript{a}yanglei.song@queensu.ca
\textsuperscript{2}Department of Statistics, University of Illinois at Urbana-Champaign, 725 S. Wright Street, Champaign, IL 61820, USA.
\textsuperscript{b}xhchen@illinois.edu
\textsuperscript{3}Department of Statistics and Data Science, Cornell University, 1194 Comstock Hall, Ithaca, NY 14853, USA.
\textsuperscript{c}kk976@cornell.edu

Principled nonparametric tests for regression curvature in $\mathbb{R}^d$ are often statistically and computationally challenging. This paper introduces the stratified incomplete local simplex (SILS) tests for joint concavity of nonparametric multiple regression. The SILS tests with suitable bootstrap calibration are shown to achieve simultaneous guarantees on dimension-free computational complexity, polynomial decay of the uniform error-in-size, and power consistency for general (global and local) alternatives. To establish these results, we develop a general theory for incomplete $U$-processes with stratified random sparse weights. Novel technical ingredients include maximal inequalities for the supremum of multiple incomplete $U$-processes.

Keywords: Nonparametric regression; curvature testing; incomplete $U$-processes; stratification

References


On algebraic Stein operators for Gaussian polynomials

EHSAN AZMOODEH¹,a, DARIO GASBARRA²,b and ROBERT E. GAUNT³,c

¹Department of Mathematical Sciences, Mathematical Sciences Bldg, The University of Liverpool, Liverpool L69 7ZL, UK. aehsan.azmoodeh@liverpool.ac.uk
²Department of Mathematics and Statistics, University of Helsinki, P.O.Box 68 00014, Finland. bdario.gasbarra@helsinki.fi
³Department of Mathematics, The University of Manchester, Oxford Road, Manchester M13 9PL, UK. crobert.gaunt@manchester.ac.uk

The first essential ingredient to build up Stein’s method for a continuous target distribution is to identify a so-called Stein operator, namely a linear differential operator with polynomial coefficients. In this paper, we introduce the notion of algebraic Stein operators (see Definition 3.4), and provide a novel algebraic method to find all the algebraic Stein operators up to a given order and polynomial degree for a target random variable of the form $Y = h(X)$, where $X = (X_1, \ldots, X_d)$ has i.i.d. standard Gaussian components and $h \in \mathbb{K}[X]$ is a polynomial with coefficients in the ring $\mathbb{K}$. Our approach links the existence of an algebraic Stein operator with null controllability of a certain linear discrete system. A MATLAB code checks the null controllability up to a given finite time $T$ (the order of the differential operator), and provides all null control sequences (polynomial coefficients of the differential operator) up to a given maximum degree $m$. This is the first paper that connects Stein’s method with computational algebra to find Stein operators for highly complex probability distributions, such as $H_p(X_1)$, where $H_p$ is the $p$-th Hermite polynomial. Some examples of Stein operators for $H_p(X_1)$, $p = 3, 4, 5, 6$, are gathered in the Appendix and many other examples are given in the Supplementary Information.

Keywords: Stein’s method; Stein operator; Gaussian integration by parts; Malliavin calculus; linear system theory; null controllability; symbolic computation; Hermite polynomials

References


Targeted cross-validation

JIAWEI ZHANG\textsuperscript{a}, JIE DING\textsuperscript{b} and YUHONG YANG\textsuperscript{c}

School of Statistics, University of Minnesota, MN 55455, USA. \textsuperscript{a}zhan4362@umn.edu, \textsuperscript{b}dingj@umn.edu, \textsuperscript{c}yangx374@umn.edu

In many applications, we have access to the complete dataset but are only interested in the prediction of a particular region of predictor variables. A standard approach is to find the globally best modeling method from a set of candidate methods. However, it is perhaps rare in reality that one candidate method is uniformly better than the others. A natural approach for this scenario is to apply a weighted $L_2$ loss in performance assessment to reflect the region-specific interest. We propose a targeted cross-validation (TCV) to select models or procedures based on a general weighted $L_2$ loss. We show that the TCV is consistent in selecting the best performing candidate under the weighted $L_2$ loss. Experimental studies are used to demonstrate the use of TCV and its potential advantage over the global CV or the approach of using only local data for modeling a local region.

Previous investigations on CV have relied on the condition that when the sample size is large enough, the ranking of two candidates stays the same. However, in many applications with the setup of changing data-generating processes or highly adaptive modeling methods, the relative performance of the methods is not static as the sample size varies. Even with a fixed data-generating process, it is possible that the ranking of two methods switches infinitely many times. In this work, we broaden the concept of the selection consistency by allowing the best candidate to switch as the sample size varies, and then establish the consistency of the TCV. This flexible framework can be applied to high-dimensional and complex machine learning scenarios where the relative performances of modeling procedures are dynamic.

Keywords: Consistency; cross-validation; model selection; regression

References


Adaptive Bayesian inference for current status data on a grid

MINWOO CHAE

Department of Industrial and Management Engineering, Pohang University of Science and Technology, 77 Cheongam-Ro, Nam-Gu, Pohang, Gyeongbuk 37673, South Korea. a mchae@postech.ac.kr

We study a Bayesian approach to the inference of an event time distribution in the current status model where observation times are supported on a grid of potentially unknown sparsity and multiple subjects share the same observation time. The model leads to a very simple likelihood, but statistical inferences are non-trivial due to the unknown sparsity of the grid. In particular, for an inference based on the maximum likelihood estimator, one needs to estimate the density of the event time distribution which is challenging because the event time is not directly observed. We consider Bayes procedures with a Dirichlet prior on the event time distribution. With this prior, the Bayes estimator and credible sets can be easily computed via a Gibbs sampler algorithm. Our main contribution is to provide thorough investigation of frequentist’s properties of the posterior distribution. Specifically, it is shown that the posterior convergence rate is adaptive to the unknown sparsity of the grid. If the grid is sufficiently sparse, we further prove the Bernstein–von Mises theorem which guarantees frequentist’s validity of Bayesian credible sets. A numerical study is also conducted for illustration.

Keywords: Adaptive procedure; Bayesian survival analysis; Bernstein–von Mises theorem; posterior convergence rate; current status model; interval-censored data

References


Bootstrapping the operator norm in high dimensions: Error estimation for covariance matrices and sketching

MILES E. LOPES¹,ᵃ, N. BENJAMIN ERICHSON²,ᵇ and MICHAEL W. MAHONEY²,ᶜ

¹Department of Statistics, University of California, Davis, Mathematical Sciences Building, Davis, CA, 95616, USA, ᵃmelopes@ucdavis.edu
²Department of Statistics, University of California, Berkeley, Evans Hall, Berkeley, CA, 94720, USA and International Computer Science Institute, Berkeley, CA, 94704, USA, ᵇerichson@berkeley.edu, ᶜmmahoney@stat.berkeley.edu

Although the operator (spectral) norm is one of the most widely used metrics for covariance estimation, comparatively little is known about the fluctuations of error in this norm. To be specific, let \( \hat{\Sigma} \) denote the sample covariance matrix of \( n \) i.i.d. observations in \( \mathbb{R}^p \) that arise from a population matrix \( \Sigma \), and let \( T_n = \sqrt{n} \| \hat{\Sigma} - \Sigma \|_{\text{op}} \). In the setting where the eigenvalues of \( \Sigma \) have a decay profile of the form \( \lambda_j(\Sigma) \approx j^{-2\beta} \), we analyze how well the bootstrap can approximate the distribution of \( T_n \). Our main result shows that up to factors of \( \log(n) \), the bootstrap can approximate the distribution of \( T_n \) with respect to the Kolmogorov metric at the rate of \( n^{-\beta^{-1}/2 + 4} \), which does not depend on the ambient dimension \( p \). In addition, we offer a supporting result of independent interest that establishes a high-probability upper bound for \( T_n \) based on flexible moment assumptions. More generally, we discuss the consequences of our work beyond covariance matrices, and show how the bootstrap can be used to estimate the errors of sketching algorithms in randomized numerical linear algebra (RandNLA). An illustration of these ideas is also provided with a climate data example.

Keywords: Bootstrap; error estimation; high-dimensional statistics; covariance estimation; randomized numerical linear algebra; sketching

References


Bootstrapping the operator norm


On consistency and sparsity for high-dimensional functional time series with application to autoregressions

SHAOJUN GUO\textsuperscript{1,a} and XINGHAO QIAO\textsuperscript{2,b}

\textsuperscript{1}Institute of Statistics and Big Data, Renmin University of China, Beijing, 100872, P.R. China. \textsuperscript{a}sjguo@ruc.edu.cn
\textsuperscript{2}Department of Statistics, London School of Economics, London, WC2A 2AE, U.K.. \textsuperscript{b}x.qiao@lse.ac.uk

Modelling a large collection of functional time series arises in a broad spectral of real applications. Under such a scenario, not only the number of functional variables can be diverging with, or even larger than the number of temporally dependent functional observations, but each function itself is an infinite-dimensional object, posing a challenging task. In this paper, we propose a three-step procedure to estimate high-dimensional functional time series models. To provide theoretical guarantees for the three-step procedure, we focus on multivariate stationary processes and propose a novel functional stability measure based on their spectral properties. Such stability measure facilitates the development of some useful concentration bounds on sample (auto)covariance functions, which serve as a fundamental tool for further convergence analysis in high-dimensional settings. As functional principal component analysis (FPCA) is one of the key dimension reduction techniques in the first step, we also investigate the non-asymptotic properties of the relevant estimated terms under a FPCA framework. To illustrate with an important application, we consider vector functional autoregressive models and develop a regularization approach to estimate autoregressive coefficient functions under the sparsity constraint. Using our derived non-asymptotic results, we investigate convergence properties of the regularized estimate under high-dimensional scaling. Finally, the finite-sample performance of the proposed method is examined through both simulations and a public financial dataset.

Keywords: Functional principal component analysis; functional stability measure; high-dimensional functional time series; non-asymptotics; sparsity; vector functional autoregression

References


Bathia, N., Yao, Q. and Ziegelmann, F. (2010). Identifying the finite dimensionality of curve time series. \textit{Ann. Statist.} \textbf{38} 3352–3386. \texttt{MR2766855} \url{https://doi.org/10.1214/10-AOS819}


We study the problem of predicting as well as the best linear predictor in a bounded Euclidean ball with respect to the squared loss. When only boundedness of the data generating distribution is assumed, we establish that the least squares estimator constrained to a bounded Euclidean ball does not attain the classical $O(d/n)$ excess risk rate, where $d$ is the dimension of the covariates and $n$ is the number of samples. In particular, we construct a bounded distribution such that the constrained least squares estimator incurs an excess risk of order $\Omega(d^3/2n)$ hence refuting a recent conjecture of Ohad Shamir [JMLR 2015]. In contrast, we observe that non-linear predictors can achieve the optimal rate $O(d/n)$ with no assumptions on the distribution of the covariates. We discuss additional distributional assumptions sufficient to guarantee an $O(d/n)$ excess risk rate for the least squares estimator. Among them are certain moment equivalence assumptions often used in the robust statistics literature. While such assumptions are central in the analysis of unbounded and heavy-tailed settings, our work indicates that in some cases, they also rule out unfavorable bounded distributions.

Keywords: Constrained least squares; ridge regression; empirical processes; average stability; Vovk-Azoury-Warmuth forecaster

References


Suboptimality of constrained least squares


Quadratic variation and quadratic roughness

RAMA CONT\(^a\) and PURBA DAS\(^b\)

**Mathematical Institute, University of Oxford, Oxford, England, \(^a\)Rama.Cont@maths.ox.ac.uk, \(^b\)Purba.Das@maths.ox.ac.uk**

We study the concept of quadratic variation of a continuous path along a sequence of partitions and its dependence with respect to the choice of the partition sequence. We introduce the concept of *quadratic roughness* of a path along a partition sequence and show that for Hölder-continuous paths satisfying this roughness condition, the quadratic variation along balanced partitions is invariant with respect to the choice of the partition sequence. Typical paths of Brownian motion are shown to satisfy this quadratic roughness property almost-surely along any partition with a required step size condition. Using these results we derive a formulation of the pathwise Föllmer-Itô calculus which is invariant with respect to the partition sequence. We also derive an invariance of local time under quadratic roughness.

**Keywords:** Quadratic variation; pathwise integration; Brownian motion; Itô calculus; local time; roughness

**References**


Quadratic roughness


Chung-type law of the iterated logarithm and exact moduli of continuity for a class of anisotropic Gaussian random fields

CHEUK YIN LEE$^{1,a}$ and YIMIN XIAO$^{2,b}$

$^1$Institut de mathématiques, École Polytechnique Fédérale de Lausanne, Station 8, CH-1015 Lausanne, Switzerland. $^a$cheuk.lee@epfl.ch
$^2$Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, United States. $^b$xiaoy@msu.edu

We establish a Chung-type law of the iterated logarithm and the exact local and uniform moduli of continuity for a large class of anisotropic Gaussian random fields with a harmonizable-type integral representation and the property of strong local nondeterminism. Compared with the existing results in the literature, our results do not require the assumption of stationary increments and provide more precise upper and lower bounds for the limiting constants. The results are applicable to the solutions of a class of linear stochastic partial differential equations driven by a fractional-colored Gaussian noise, including the stochastic heat equation.

Keywords: Gaussian random fields; harmonizable representation; strong local nondeterminism; law of the iterated logarithm; modulus of continuity; stochastic heat equation

References


Multilevel bootstrap particle filter

KARI HEINE\textsuperscript{a} and DANIEL BURROWS\textsuperscript{b}

Department of Mathematical Sciences, University of Bath, Bath, UK,\textsuperscript{a}k.m.p.heine@bath.ac.uk, \textsuperscript{b}dwb26@bath.ac.uk

We consider situations where the applicability of sequential Monte Carlo particle filters is compromised due to the expensive evaluation of the particle weights. To alleviate this problem, we propose a new particle filter algorithm based on the multilevel approach. We show that the resulting multilevel bootstrap particle filter (MLBPF) retains the strong law of large numbers as well as the central limit theorem of classical particle filters under mild conditions. Our numerical experiments demonstrate up to 85% reduction in computation time compared to the classical bootstrap particle filter, in certain settings. While it should be acknowledged that this reduction is highly application dependent, and a similar gain should not be expected for all applications across the board, we believe that this substantial improvement in certain settings makes MLBPF an important addition to the family of sequential Monte Carlo methods.

Keywords: Sequential Monte Carlo; particle filter; multilevel; hidden Markov model

References


Exponential and strong ergodicity for one-dimensional time-changed symmetric stable processes

TAO WANG

School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, Beijing, China. wang_tao@mail.bnu.edu.cn

We obtain explicit criteria for both exponential ergodicity and strong ergodicity for one-dimensional time-changed symmetric stable processes with \( \alpha \in (1, 2) \). Explicit lower bounds for ergodic convergence rates are given.

Keywords: Stable process; time change; Dirichlet eigenvalue; strong ergodicity; exponential ergodicity; Green operator

References

Ergodicity for time-changed stable process


A necessary and sufficient condition for the convergence of the derivative martingale in a branching Lévy process

BASTIEN MALLEIN\textsuperscript{1,a} and QUAN SHI\textsuperscript{2,b}

\textsuperscript{1}Université Sorbonne Paris Nord, LAGA, UMR 7539, F-93430, Villetaneuse, France. 
\textsuperscript{a}mallein@math.univ-paris13.fr
\textsuperscript{2}Academy of Mathematics and Systems Science, Chinese Academy of Sciences, 100190, Beijing, China. 
\textsuperscript{b}quan.shi@amss.ac.cn

A continuous-time particle system on the real line satisfying the branching property and an exponential integrability condition is called a branching Lévy process, and its law is characterized by a triplet $(\sigma^2, a, \Lambda)$. We obtain a necessary and sufficient condition for the convergence of the derivative martingale of such a process to a non-trivial limit in terms of $(\sigma^2, a, \Lambda)$. This extends previously known results on branching Brownian motions and branching random walks. To obtain this result, we rely on the spinal decomposition and establish a novel zero-one law on the perpetual integrals of centred Lévy processes conditioned to stay positive.

Keywords: Branching Lévy process; derivative martingale; spinal decomposition; Lévy process; perpetual integral

References

B. Mallein and Q. Shi


The derivative martingale in a branching Lévy process


Almost-sure asymptotics for Riemannian random waves

LOUIS GASS

Department, Univ Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France. a louis.gass@ens-rennes.fr

We consider the Riemannian random wave model of Gaussian linear combinations of Laplace eigenfunctions on a general compact Riemannian manifold. With probability one with respect to the Gaussian coefficients, we establish that, both for large and small band models, the process properly rescaled and evaluated at an independently and uniformly chosen point \( X \) on the manifold, converges in distribution under the sole randomness of \( X \) towards an universal Gaussian field as the frequency tends to infinity. This result extends the celebrated central limit Theorem of Salem–Zygmund for trigonometric polynomials series to the more general framework of compact Riemannian manifolds. We then deduce from the above convergence the almost-sure asymptotics of the nodal volume associated with the random wave. To the best of our knowledge, in the real Riemannian case, these asymptotics were only known in expectation and not in the almost sure sense due to the lack of sufficiently accurate variance estimates. This in particular addresses a question of S. Zelditch regarding the almost sure equidistribution of nodal volume.

Keywords: Geometric probability; mathematical physics

References


We study randomized designs that minimize the asymptotic variance of a debiased lasso estimator when a large pool of unlabeled data is available but measuring the corresponding responses is costly. The optimal sampling distribution arises as the solution of a semidefinite program. The improvements in efficiency that result from these optimal designs are demonstrated via simulation experiments.

Keywords: Optimal design; inference; sparsity; compressed sensing

References


High-dimensional c-optimal designs


Multivariate time series models for mixed data

ZINSOU-MAX DEBALY\(^a\) and LIONEL TRUQUET\(^b\)

\(^a\)ENSAI-CREST UMR 9194, Campus de Ker-Lann, Rue Blaise Pascal – BP 37203, 35712 BRUZ cedex, France, \(^b\)zinsou-max.debaly@ensai.fr, \(b\)lionel.truquet@ensai.fr

We introduce a general approach which unifies some previous attempts for modeling the dynamic of multivariate time series or for regression analysis when the data are of mixed type (binary/count/continuous). Our approach is quite flexible since conditionally on past values, each coordinate at time \(t\) can have a distribution compatible with a standard univariate time series model such as GARCH, ARMA, INGARCH or logistic models whereas past values of the other coordinates play the role of exogenous covariates in the dynamic. The simultaneous dependence in the multivariate time series can be modeled with a copula. Additional exogenous covariates are also allowed in the dynamic. We first study some usual stability properties of these models and then show that autoregressive parameters can be consistently estimated equation-by-equation using a pseudo-maximum likelihood method, leading to a fast implementation even when the number of time series is large. Moreover, we prove consistency results when a parametric copula model is fitted to the time series and in the case of Gaussian copulas, we show that the likelihood estimator of the correlation matrix is strongly consistent. We carefully check all our assumptions for two prototypical examples: a GARCH/INGARCH model and logistic/log-linear INGARCH model. Our results are illustrated with numerical experiments as well as two real data sets.

**Keywords:** Time series; mixed data; observation-driven models

**References**


Bootstrap percolation on the stochastic block model

GIOVANNI LUCA TORRISI1,a, MICHELE GARETTO2,b and EMILIO LEONARDI3,c

1CNR-IAC, Roma, Italy. a giovanniluca.torrisi@cnr.it
2Università di Torino, Torino, Italy. bmichele.garetto@unito.it
3Politecnico di Torino, Torino, Italy. c emilio.leonardi@polito.it

We analyze the bootstrap percolation process on the stochastic block model (SBM), a natural extension of the Erdős–Rényi random graph that incorporates the community structure observed in many real systems. In the SBM, nodes are partitioned into two subsets, which represent different communities, and pairs of nodes are independently connected with a probability that depends on the communities they belong to. Under mild assumptions on the system parameters, we prove the existence of a sharp phase transition for the final number of active nodes and characterize the sub-critical and the super-critical regimes in terms of the number of initially active nodes, which are selected uniformly at random in each community.

Keywords: Bootstrap percolation; random graphs; stochastic block model

References

Bootstrap percolation on the SBM


A trajectorial approach to relative entropy dissipation of McKean–Vlasov diffusions: Gradient flows and HWBI inequalities

BERTRAM TSCHIDERER¹,a and LANE CHUN YEUNG²,b

¹Faculty of Mathematics, University of Vienna, Austria, abertram.tschiderer@univie.ac.at
²Department of Industrial Engineering and Operations Research, Columbia University, New York, USA, blyeung@columbia.edu

We formulate a trajectorial version of the relative entropy dissipation identity for McKean–Vlasov diffusions, extending recent results which apply to non-interacting diffusions. Our stochastic analysis approach is based on time-reversal of diffusions and Lions’ differential calculus over Wasserstein space. It allows us to compute explicitly the rate of relative entropy dissipation along every trajectory of the underlying diffusion via the semimartingale decomposition of the corresponding relative entropy process. As a first application, we obtain a new interpretation of the gradient flow structure for the granular media equation, generalizing a formulation developed recently for the linear Fokker–Planck equation. Secondly, we show how the trajectorial approach leads to a new derivation of the HWBI inequality, which relates relative entropy (H), Wasserstein distance (W), barycenter (B) and Fisher information (I).

Keywords: Relative entropy dissipation; gradient flow; McKean–Vlasov diffusion; granular media equation; HWBI inequality

References


SDEs with critical time dependent drifts: Weak solutions

MICHAEL RÖCKNER\textsuperscript{1,2,a} and GUOHUAN ZHAO\textsuperscript{3,b}

\textsuperscript{1}Department of Mathematics, Bielefeld University, Germany
\textsuperscript{2}Academy of Mathematics and Systems Science, Chinese Academy of Sciences (CAS), Beijing, P.R.China,
\textsuperscript{a}roeckner@math.uni-bielefeld.de
\textsuperscript{3}Department of Mathematics, Bielefeld University, Germany, \textsuperscript{b}zhaoguohuan@gmail.com

For $d \geq 3$, we prove that time-inhomogeneous stochastic differential equations driven by additive noises with drifts in critical Lebesgue space $L^q([0,T];L^p(\mathbb{R}^d))$, where $(p,q) \in (d,\infty] \times [2,\infty)$ and $d/p + 2/q = 1$, or $(p,q) = (d,\infty)$ and $\text{div}\, b \in L^\infty([0,T];L^{d/2+\varepsilon}(\mathbb{R}^d))$, are well-posed. The weak uniqueness is obtained by solving corresponding Kolmogorov backward equations in some second-order Sobolev spaces, which is analytically interesting in itself.

Keywords: Weak solutions; Ladyzhenskaya–Prodi–Serrin condition; Kolmogorov equations; De Giorgi’s method

References


Linear multifractional stable sheets in the broad sense: Existence and joint continuity of local times

YUJIA DING$^{1,a}$, QIDI PENG$^{1,b}$ and YIMIN XIAO$^{2,c}$

$^1$Institute of Mathematical Sciences, Claremont Graduate University, USA, $^a$yujia.ding@cgu.edu, $^b$qidi.peng@cgu.edu
$^2$Department of Statistics and Probability, Michigan State University, USA, $^c$xiaoy@msu.edu

We introduce the notion of linear multifractional stable sheets in the broad sense (LMSS) with $\alpha \in (0, 2]$, to include both linear multifractional Brownian sheets ($\alpha = 2$) and linear multifractional stable sheets ($\alpha < 2$). The purpose of the present paper is to study the existence and joint continuity of the local times of LMSS, and also the local Hölder condition of the local times in the set variable. Among the main results of this paper, Theorem 2.4 provides a sufficient and necessary condition for the existence of local times of LMSS; Theorem 3.1 shows a sufficient condition for the joint continuity of local times; and Theorem 4.1 proves a sharp local Hölder condition for the local times in the set variable. All these theorems improve significantly the existing results for the local times of multifractional Brownian sheets and linear multifractional stable sheets in the literature.

Keywords: Linear multifractional Brownian sheets; linear multifractional stable sheets; local times; joint continuity

References


Non-homogeneous Poisson process intensity modeling and estimation using measure transport

TIN LOK JAMES NG\textsuperscript{a} and ANDREW ZAMMIT-MANGION\textsuperscript{b}

School of Computer Science and Statistics, Trinity College Dublin, Ireland, \textsuperscript{a}ngja@tcd.ie
School of Mathematics and Applied Statistics, University of Wollongong, Australia, \textsuperscript{b}azm@uow.edu.au

Non-homogeneous Poisson processes are used in a wide range of scientific disciplines, ranging from the environmental sciences to the health sciences. Often, the central object of interest in a point process is the underlying intensity function. Here, we present a general model for the intensity function of a non-homogeneous Poisson process using measure transport. The model is built from a flexible bijective mapping that maps from the underlying intensity function of interest to a simpler reference intensity function. We enforce bijectivity by modeling the map as a composition of multiple bijective maps that have increasing triangular structure, and show that the model exhibits an important approximation property. Estimation of the flexible mapping is accomplished within an optimization framework, wherein computations are efficiently done using tools originally designed to facilitate deep learning, and a graphics processing unit. Point process simulation and uncertainty quantification are straightforward to do with the proposed model. We demonstrate the potential benefits of our proposed method over conventional approaches to intensity modeling through various simulation studies. We also illustrate the use of our model on a real data set containing the locations of seismic events near Fiji since 1964.

Keywords: Poisson point process; intensity estimation; measure transport; deep neural network

References


General construction and classes of explicit \(L^1\)-optimal couplings

GIOVANNI PUCETTI\(^1\),\(^a\) and LUDGER RÜSCHENDORF\(^2\),\(^b\)

\(^1\)University of Milano, Via Conservatorio 7, 20122 Milano, Italy, \(^a\)giovanni.puccetti@unimi.it
\(^2\)University of Freiburg, Ernst-Zermelo-Straße 1, 79104 Freiburg, Germany, \(^b\)ruschen@stochastik.uni-freiburg.de

The main scope of this paper is to give some explicit classes of examples of \(L^1\)-optimal couplings. Optimal transport w.r.t. the Kantorovich metric \(\ell_1\) (resp. the Wasserstein metric \(W_1\)) between two absolutely continuous measures is known since the basic papers of Kantorovich and Rubinstein (\textit{Dokl. Akad. Nauk SSSR} 115 (1957) 1058–1061) and Sudakov (\textit{Proc. Steklov Inst. Math.} 141 (1979) 1–178) to occur on rays induced by a decomposition of the basic space (and more generally to higher dimensional decompositions in the case of general measures) induced by the corresponding dual potentials. Several papers have given this kind of structural result and established existence and uniqueness of solutions in varying generality. Since the dual problems pose typically too strong challenges to be solved in explicit form, these structural results have so far been applied for the solution of few particular instances.

First, we give a self-contained review of some basic optimal coupling results and we propose and investigate in particular some basic principles for the construction of \(L^1\)-optimal couplings given by a reduction principle and some usable forms of the decomposition method. This reduction principle, together with symmetry properties of the reduced measures, gives a hint to the decomposition of the space into sectors and via the non crossing property of optimal transport leads to the choice of transportation rays. The optimality of the induced transports is then a consequence of the characterization results of optimal couplings.

Then, we apply these principles to determine in explicit form \(L^1\)-optimal couplings for several classes of examples of elliptical distributions. In particular, we give for the first time a general construction of \(L^1\)-optimal couplings between two bivariate Gaussian distributions. We also discuss optimality of special constructions like shifts and scalings, and provide an extended class of dual functionals allowing for the closed-form computation of the \(\ell_1\)-metric or of accurate lower bounds of it in a variety of examples.

\textbf{Keywords:} Kantorovich \(\ell_1\)-metric; \(L^1\)-Wasserstein distance; optimal mass transportation; optimal couplings; Gaussian distributions; Monge-Kantorovich problem; Kantorovich-Rubinstein Theorem

\textbf{References}


$L^1$-optimal couplings


