CONTENTS

DEL BARRIO, E., CUESTA-ALBERTOS, J.A. and MATRAN, C. 2123
The complex behaviour of Galton rank-order statistic

BELOMESTNY, D., GUGUSHVILI, S., SCHAUER, M. and SPREIJ, P. 2151
Nonparametric Bayesian volatility estimation for gamma-driven stochastic differential equations

BARP, A., OATES, C.J., PORCU, E. and GIROLAMI, M. 2181
A Riemann–Stein kernel method

BETKEN, A. and WENDLER, M. 2209
Rank-based change-point analysis for long-range dependent time series

MALLER, R., RESNICK, S. and SHEMEHSAVAR, S. 2234
Splitting the sample at the largest uncensored observation

VAN DELFT, A. and DETTE, H. 2260
Pivotal tests for relevant differences in the second order dynamics of functional time series

CATTIAUX, P. and GUILLIN, A. 2294
Functional inequalities for perturbed measures with applications to log-concave measures and to some Bayesian problems

GONG, F. and SUN, X. 2322
On the characterization of Brownian bridge measure on the pinned path space over a compact Riemannian manifold

GAMBOA, F., GREMAUD, P., KLEIN, T. and LAGNOUX, A. 2345
Global sensitivity analysis: A novel generation of mighty estimators based on rank statistics

BAI, Z., CHOI, K.P., FUJIKOSHI, Y. and HU, J. 2375
Asymptotics of AIC, BIC and $C_p$ model selection rules in high-dimensional regression

BERTAZZI, A. and BIERKENS, J. 2404
Adaptive schemes for piecewise deterministic Monte Carlo algorithms

BRAS, P. 2431
Convergence rates of Gibbs measures with degenerate minimum

YUAN, M. and SHANG, Z. 2459
Sharp detection boundaries on testing dense subhypergraph

RANDRIANARISOA, T. 2492
Smoothing and adaptation of shifted Pólya tree ensembles

BACHOC, F., PORCU, E., BEVILACQUA, M., FURRER, R. and FAOUZI, T. 2518
Asymptotically equivalent prediction in multivariate geostatistics

COSTA, M. and JORDAN, J. 2546
Phase transitions in non-linear urns with interacting types

(continued)

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

(continued)

JANSON, S. and THACKER, D.
Continuous-time digital search tree and a border aggregation model 2563

LODHIA, A., LEVIN, K. and LEVINA, E.
Matrix means and a novel high-dimensional shrinkage phenomenon 2578

GLOAGUEN, P., LE CORFF, S. and OLSSON, J.
A pseudo-marginal sequential Monte Carlo online smoothing algorithm 2606

JIANG, H., WAN, Y. and YANG, G.
Deviation inequalities and Cramér-type moderate deviations for the explosive autoregressive process 2634

NILES-WEED, J. and RIGOLLET, P.
Estimation of Wasserstein distances in the Spiked Transport Model 2663

MUSTA, E., PATILEA, V. and VAN KEILEGOM, I.
A presmoothing approach for estimation in the semiparametric Cox mixture cure model 2689

FONTES, L.R., FRÓMETA, S. and ZUAZNÁBAR, L.
Asymptotic behavior of a low-temperature non-cascading 2-GREM dynamics at extreme time scales 2716

CARPENTIER, A., COLLIER, O., COMMINGES, L., TSYBAKOV, A.B. and WANG, Y.
Estimation of the $\ell_2$-norm and testing in sparse linear regression with unknown variance 2744

CUI, T. and TONG, X.T.
A unified performance analysis of likelihood-informed subspace methods 2788

WIESEL, J.C.W.
Measuring association with Wasserstein distances 2816

PILIPAUSKAITĖ, V. and SURGAILIS, D.
Local scaling limits of Lévy driven fractional random fields 2833

ORENSTEIN, P.
Robust importance sampling with adaptive winsorization 2862

KÜHN, F.
Upper functions for sample paths of Lévy(-type) processes 2874

DETTE, H., DIERICKX, G. and KUTTA, T.
Quantifying deviations from separability in space-time functional processes 2909

JUN, W., JIAHUI, X., LONG, Y. and WANG, Z.
Tracy-Widom limit for the largest eigenvalue of high-dimensional covariance matrices in elliptical distributions 2941

LIU, L., SU, W., YIN, G., ZHAO, X. and ZHANG, Y.
Nonparametric inference for reversed mean models with panel count data 2968

CALONICO, S., CATTANEO, M.D. and FARRELL, M.H.
Coverage error optimal confidence intervals for local polynomial regression 2998

MOGENSEN, S.W. and HANSEN, N.R.
Graphical modeling of stochastic processes driven by correlated noise 3023

LANTERI, A., MAGGIONI, M. and VIGOGNA, S.
Conditional regression for single-index models 3051
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: http://www.bernoulli-society.org/index.php/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 http://www.bernoulli-society.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.

©2022 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 2022 Bernoulli consists of 4 issues published in February, May, August and November.
The complex behaviour of Galton rank-order statistic

EUSTASIO DEL BARRIO¹,a, JUAN A. CUESTA-ALBERTOS²,c and CARLOS MATRAN¹,b

¹Department of Statistics and Operational Research and IMUVA, University of Valladolid, Valladolid, Spain.
aeustasio.delbarrio@uva.es, b carlos.matran@uva.es
²Departmento de Matemáticas, Estadística y Computación, Universidad de Cantabria, Santander, Spain.
cuestaj@unican.es

Galton’s rank-order statistic is one of the oldest statistical tools for two-sample comparisons. It is also a very natural index to measure departures from stochastic dominance. Yet, its asymptotic behaviour has been investigated only partially, under restrictive assumptions. This work provides a comprehensive study of this behaviour, based on the analysis of the so-called contact set (a modification of the set in which the quantile functions coincide). We show that a.s. convergence to the population counterpart holds if and only if the contact set has zero Lebesgue measure. When this set is finite we show that the asymptotic behaviour is determined by the local behaviour of a suitable reparameterization of the quantile functions in a neighbourhood of the contact points. Regular crossings result in standard rates and Gaussian limiting distributions, but higher order contacts (in the sense introduced in this work) or contacts at the extremes of the supports may result in different rates and non-Gaussian limits.

Keywords: Relaxed stochastic dominance; asymptotics; consistency; Galton rank-order statistic; comparison of quantile functions; contact points; crossings; tangencies; contact intensity

References


Nonparametric Bayesian volatility estimation for gamma-driven stochastic differential equations

DENIS BELOMESTNY\textsuperscript{1,2,a}, SHOTA GUGUSHVILI\textsuperscript{3,b}, MORITZ SCHAUER\textsuperscript{4,c} and PETER SPREIJ\textsuperscript{5,6,d}

\textsuperscript{1}Faculty of Mathematics, Duisburg-Essen University, Thea-Leymann-Str. 9, D-45127 Essen, Germany
\textsuperscript{2}Faculty of Computer Sciences, HSE University. \textsuperscript{a}denis.belomestny@uni-due.de
\textsuperscript{3}Wageningen University & Research, Postbus 16, 6700 AA Wageningen, The Netherlands. \textsuperscript{b}shota@yesdatasolutions.com
\textsuperscript{4}Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE-412 96 Göteborg, Sweden. \textsuperscript{c}smoritz@chalmers.se
\textsuperscript{5}Korteweg-de Vries Institute for Mathematics, Universiteit van Amsterdam, P.O. Box 94248, 1090 GE Amsterdam, The Netherlands
\textsuperscript{6}Institute for Mathematics, Astrophysics and Particle Physics, Radboud University, Nijmegen, Netherlands. \textsuperscript{d}spreij@uva.nl

We study a nonparametric Bayesian approach to estimation of the volatility function of a stochastic differential equation driven by a gamma process. The volatility function is modelled a priori as piecewise constant, and we specify a gamma prior on its values. This leads to a straightforward procedure for posterior inference via an MCMC procedure. We give theoretical performance guarantees (minimax optimal contraction rates for the posterior) for the Bayesian estimate in terms of the regularity of the unknown volatility function. We illustrate the method on synthetic and real data examples.

Keywords: Gamma process; nonparametric Bayesian estimation; stochastic differential equation

References


A Riemann–Stein kernel method

ALESSANDRO BARP1,2,a, CHRIS. J. OATES3,2,c, EMILIO PORCU4,5,d and MARK GIROLAMI1,2,b

1University of Cambridge, Cambridge, UK. aab2286@eng.cam.ac.uk, bmag92@eng.cam.ac.uk
2Alan Turing Institute, London, UK
3Newcastle University, Newcastle-upon-Tyne, UK. cchris.oates@ncl.ac.uk
4Khalifa University, Abu Dhabi, United Arab Emirates. demilio.porcu@ku.ac.ae
5Trinity College Dublin, Dublin, IE

This paper proposes and studies a numerical method for approximation of posterior expectations based on interpolation with a Stein reproducing kernel. Finite-sample-size bounds on the approximation error are established for posterior distributions supported on a compact Riemannian manifold, and we relate these to a kernel Stein discrepancy (KSD). Moreover, we prove in our setting that the KSD is equivalent to Sobolev discrepancy and, in doing so, we completely characterise the convergence-determining properties of KSD. Our contribution is rooted in a novel combination of Stein’s method, the theory of reproducing kernels, and existence and regularity results for partial differential equations on a Riemannian manifold.

Keywords: Bayesian computation; kernel Stein discrepancy; reproducing kernel; Stein equation

References


A Riemann–Stein kernel method


We consider change-point tests based on rank statistics to test for structural changes in long-range dependent observations. Under the hypothesis of stationary time series and under the assumption of a change with decreasing change-point height, the asymptotic distributions of corresponding test statistics are derived. For this, a uniform reduction principle for the sequential empirical process in a two-parameter Skorohod space equipped with a weighted supremum norm is proved. Moreover, we compare the efficiency of rank tests resulting from the consideration of different score functions. Under Gaussianity, the asymptotic relative efficiency of rank-based tests with respect to the CuSum test is 1, irrespective of the score function. Regarding the practical implementation of rank-based change-point tests, we suggest to combine self-normalized rank statistics with subsampling. The theoretical results are accompanied by simulation studies that, in particular, allow for a comparison of rank tests resulting from different score functions. With respect to the finite sample performance of rank-based change-point tests, the Van der Waerden rank test proves to be favorable in a broad range of situations. Finally, we analyze data sets from economy, hydrology, and network traffic monitoring in view of structural changes and compare our results to previous analysis of the data.

Keywords: Rank statistic; change-point; long memory; self-normalization; subsampling; empirical process; asymptotic relative efficiency

References


Splitting the sample at the largest uncensored observation

ROSS MALLER1,a, SIDNEY RESNICK2,b and SOUDABEH SHEMEHSAVAR3,c

1 Research School of Finance, Actuarial Studies & Statistics, The Australian National University, Canberra, Australia. a Ross.Maller@anu.edu.au
2 School of Operations Research & Information Engineering, Cornell University, Ithaca, New York, USA. b sir1@cornell.edu
3 School of Mathematics, Statistics and Computer Science, University of Tehran, Tehran, Iran. c shemehsavar@ut.ac.ir

We calculate finite sample and asymptotic distributions for the largest censored and uncensored survival times, and some related statistics, from a sample of survival data generated according to an iid censoring model. These statistics are important for assessing whether there is sufficient follow-up in the sample to be confident of the presence of immune or cured individuals in the population. A key structural result obtained is that, conditional on the value of the largest uncensored survival time, and knowing the number of censored observations exceeding this time, the sample partitions into two independent subsamples, each subsample having the distribution of an iid sample of censored survival times, of reduced size, from truncated random variables. This result provides valuable insight into the construction of censored survival data, and facilitates the calculation of explicit finite sample formulae. We illustrate by calculating distributions of statistics useful for testing for sufficient follow-up in a sample, and apply extreme value methods to derive asymptotic distributions for some of those.

Keywords: Survival data; largest censored and uncensored survival times; sufficient follow-up; immune or cured individuals; cure model; extreme value methods

References


Pivotal tests for relevant differences in the second order dynamics of functional time series

ANNE VAN DELFT\textsuperscript{1,\textdagger} and HOLGER DETTE\textsuperscript{2,\textdagger}

\textsuperscript{1}Department of Statistics, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027, USA.
\textsuperscript{a}anne.vandelft@columbia.edu
\textsuperscript{2}Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany. \textsuperscript{b}holger.dette@rub.de

Motivated by the need to statistically quantify differences between modern (complex) data-sets which commonly result as high-resolution measurements of stochastic processes varying over a continuum, we propose novel testing procedures to detect relevant differences between the second order dynamics of two functional time series. In order to take the between-function dynamics into account that characterize this type of functional data, a frequency domain approach is taken. Test statistics are developed to compare differences in the spectral density operators and in the primary modes of variation as encoded in the associated eigenelements. Under mild moment conditions, we show convergence of the underlying statistics to Brownian motions and construct pivotal test statistics. The latter is essential because the nuisance parameters can be unwieldy and their robust estimation infeasible, especially if the two functional time series are dependent. In addition to these novel features, the properties of the tests are robust to any choice of frequency band enabling also to compare energy contents at a single frequency. The finite sample performance of the tests are verified through a simulation study and are illustrated with an application to fMRI data.

Keywords: Functional data; time series; spectral analysis; relevant tests; self-normalization; martingale theory

References


Pivotal tests for relevant differences in the second order dynamics of functional time series


We study functional inequalities (Poincaré, Cheeger, log-Sobolev) for probability measures obtained as perturbations. Several explicit results for general measures as well as log-concave distributions are given. The initial goal of this work was to obtain explicit bounds on the constants in view of statistical applications. These results are then applied to the Langevin Monte-Carlo method used in statistics in order to compute Bayesian estimators.

**Keywords:** Logconcave measure; Poincaré inequality; Cheeger inequality; logarithmic Sobolev inequality; perturbation; bayesian statistic; sparse learning

**References**


Inequalities for perturbed measures


On the characterization of Brownian bridge measure on the pinned path space over a compact Riemannian manifold

FUZHOU GONG$^{1,2,a}$ and XIAOXIA SUN$^{3,b}$

$^1$Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China
$^2$School of Mathematical Science, University of the Chinese Academy of Sciences, Beijing, China.
$^a$fzgong@amt.ac.cn
$^3$School of Data Science and Artificial Intelligence, Dongbei University of Finance and Economics, Dalian, China. $^b$xiaoxiasun@dufe.edu.cn

In this paper, we focus on the characterization of a Brownian bridge measure on the pinned path space over a compact Riemannian manifold. In the case when the Riemannian manifold is simply connected, we prove that the integration by parts formula can characterize the Brownian bridge measure. Otherwise, we show that it is not always true by constructing an illustrating example.

Keywords: Brownian bridge; pull back formula; integration by parts formula; characterization

References


Characterization of Brownian bridge measure


Global sensitivity analysis: A novel generation of mighty estimators based on rank statistics

FABRICE GAMBOA¹,a, PIERRE GREMAUD²,b, THIERRY KLEIN³,c and AGNÈS LAGNOUX⁴,d

¹Institut de Mathématiques de Toulouse and ANITI; UMR5219, Université de Toulouse; CNRS, UT3, F-31062 Toulouse, France. a fabrice.gamboa@math.univ-toulouse.fr
²Department of Mathematics, NC State University, Raleigh, North Carolina 27695, USA. b gremaud@ncsu.edu
³Institut de Mathématiques de Toulouse; UMR5219, Université de Toulouse; ENAC - Ecole Nationale de l’Aviation Civile, Université de Toulouse, France. c thierry01.klein@enac.fr
⁴Institut de Mathématiques de Toulouse; UMR5219, Université de Toulouse; CNRS, UT2J, F-31058 Toulouse, France. d lagnoux@univ-tlse2.fr

We propose a new statistical estimation framework for a large family of global sensitivity analysis indices. Our approach is based on rank statistics and uses an empirical correlation coefficient recently introduced by Chatterjee (Calcutta Statist. Assoc. Bull. 33 (1984) 1–2). We show how to apply this approach to compute not only the Cramér-von-Mises indices, directly related to Chatterjee’s notion of correlation, but also first-order Sobol’ indices, general metric space indices and higher-order moment indices. We establish consistency of the resulting estimators and demonstrate their numerical efficiency, especially for small sample sizes. In addition, we prove a central limit theorem for the estimators of the first-order Sobol’ indices.

Keywords: Global sensitivity analysis; Cramér-von-Mises distance; Pick-Freeze method; Chatterjee’s coefficient of correlation; Sobol’ indices estimation

References


Asymptotics of AIC, BIC and $C_p$ model selection rules in high-dimensional regression

ZHIDONG BAI$^{1,a}$, KWOK PUI CHOI$^{2,c}$, YASUNORI FUJIKOSHI$^{3,d}$ and JIANG HU$^{1,b}$

$^1$KLASMOE and School of Mathematics & Statistics, Northeast Normal University, China. $^a$baizd@nenu.educn, $^b$huj156@nenu.educn
$^2$Department of Statistics and Data Science, National University of Singapore, Singapore. $^c$stackp@nus.edusg
$^3$Department of Mathematics, Graduate School of Science, Hiroshima University, Japan. $^d$fujikoshi_y@yahoo.co.jp

Variable selection in multivariate linear regression is essential for the interpretation, subsequent statistical inferences and predictions of the statistical problem at hand. It has a long history of being studied, and many regressor selection criteria have been proposed. Most commonly used criteria include the Akaike information criterion (AIC), Bayesian information criterion (BIC), and Mallow’s $C_p$ and their modifications. It is well-known that if the true model is among the candidate models, then BIC is strongly consistent while AIC is not when only the sample size tends to infinity and the numbers of response variables and regressors remain fixed; a setting often described as large-sample. Increasingly, more and more datasets are viewed as high-dimensional in the sense that the number of response variables ($p$), the number of regressors ($k$) and the sample size ($n$) tend to infinity such that $p/n \to c \in (0,1)$ and $k/n \to \alpha \in [0,1)$ with $\alpha + c < 1$. A few recent works reported that, under high dimension, the asymptotic properties of AIC, BIC and $C_p$ selection rules in the large-sample setting do not necessarily carry over in the high-dimensional setting. In this paper, we clarify their asymptotic properties and provide sufficient conditions for which a selection rule is strongly consistent, almost surely under specify and over specify a true model. We do not assume normality in the errors, and we only require finite fourth moment. The main tool employed is random matrix theory techniques. A consequence of this work states that, under certain mild high-dimensional conditions, if the BIC selection rule is strongly consistent then the AIC selection rule is also strongly consistent, but not vice versa. This result is in stark contrast to the large-sample result.

Keywords: AIC; BIC; $C_p$; strong consistency; high-dimensional criteria; multi-response regression; variable selection; RMT

References


Adaptive schemes for piecewise deterministic Monte Carlo algorithms

ANDREA BERTAZZI\textsuperscript{a} and JORIS BIERKENS\textsuperscript{b}

\textit{Delft Institute of Applied Mathematics, TU Delft, Mekelweg 4, 2628 CD Delft, The Netherlands.}
\textsuperscript{a}a.bertazzi@tudelft.nl, \textsuperscript{b}joris.bierkens@tudelft.nl

The Bouncy Particle sampler (BPS) and the Zig-Zag sampler (ZZS) are continuous time, non-reversible Monte Carlo methods based on piecewise deterministic Markov processes. Experiments show that the speed of convergence of these samplers can be affected by the shape of the target distribution, as for instance in the case of anisotropic targets. We propose an adaptive scheme that iteratively learns all or part of the covariance matrix of the target and takes advantage of the obtained information to modify the underlying process with the aim of increasing the speed of convergence. Moreover, we define an adaptive scheme that automatically tunes the refreshment rate of the BPS or ZZS. We prove ergodicity and a law of large numbers for all the proposed adaptive algorithms. Finally, we show the benefits of the adaptive samplers with several numerical simulations.

\textit{Keywords:} Adaptive Markov process Monte Carlo; piecewise deterministic Markov processes; bouncy particle sampler; zig-zag sampler; ergodicity

\textbf{References}


Adaptive schemes for PDMC algorithms


Convergence rates of Gibbs measures with degenerate minimum

PIERRE BRAS

Laboratoire de Probabilités, Statistique et Modélisation, UMR 8001, case 188, Sorbonne Université, 4 pl. Jussieu, F-75252 Paris Cedex 5, France. a pierre.bras@sorbonne-universite.fr

We study convergence rates of Gibbs measures, with density proportional to \( \exp(-f(x)/t) \), as \( t \to 0 \) where \( f : \mathbb{R}^d \to \mathbb{R} \) admits a unique global minimum at \( x^* \). We focus on the case where the Hessian is not definite at \( x^* \). We assume instead that the minimum is strictly polynomial and give a higher order nested expansion of \( f \) at \( x^* \), which depends on every coordinate. We give an algorithm yielding such an expansion if the polynomial order of \( x^* \) is no more than 8, in connection with Hilbert’s 17th problem. However, we prove that the case where the order is 10 or higher is fundamentally different and that further assumptions are needed. We then give the rate of convergence of Gibbs measures using this expansion. Finally we adapt our results to the multiple well case.

Keywords: Gibbs measures; stochastic optimization; central limit theorem; Hilbert’s 17th problem

References


Sharp detection boundaries on testing dense subhypergraph

MINGAO YUAN\textsuperscript{1,a} and ZUOFENG SHANG\textsuperscript{2,b}

\textsuperscript{1}Department of Statistics, North Dakota State University, 1340 Administration Ave, Fargo, ND 58105, USA. \textsuperscript{a}mingao.yuan@ndsu.edu

\textsuperscript{2}Department of Mathematical Sciences, New Jersey Institute of Technology, 323 Dr Martin Luther King Jr Blvd, Newark, NJ 07102, USA. \textsuperscript{b}zshang@njit.edu

We study the problem of testing the existence of a dense subhypergraph. The null hypothesis corresponds to an Erdős-Rényi uniform random hypergraph and the alternative hypothesis corresponds to a uniform random hypergraph that contains a dense subhypergraph. We establish sharp detection boundaries in both scenarios: (1) the edge probabilities are known; (2) the edge probabilities are unknown. In both scenarios, sharp detection boundaries are characterized by the appropriate model parameters. Asymptotically powerful tests are provided when the model parameters fall in the detectable regions. Our results indicate that the detectable regions for general hypergraph models are dramatically different from their graph counterparts.

Keywords: Sharp detection boundary; uniform hypergraph; dense subhypergraph detection; asymptotically powerful test

References


Smoothing and adaptation of shifted Pólya tree ensembles

THIBAULT RANDRIANARISOA

LPSM, Sorbonne Université, Paris, France. thibault.randrianarisoa@sorbonne-universite.fr

Recently, S. Arlot and R. Genuer have shown that a random forest model outperforms its single-tree counterpart in estimating $\alpha$-Hölder functions, $1 \leq \alpha \leq 2$. This backs up the idea that ensembles of tree estimators are smoother estimators than single trees. On the other hand, most positive optimality results on Bayesian tree-based methods assume that $\alpha \leq 1$. Naturally, one wonders whether Bayesian counterparts of forest estimators are optimal on smoother classes, just as observed with frequentist estimators for $\alpha \leq 2$. We focus on density estimation and introduce an ensemble estimator from the classical (truncated) Pólya tree construction in Bayesian nonparametrics. Inspired by the work mentioned above, the resulting Bayesian forest estimator is shown to lead to optimal posterior contraction rates, up to logarithmic terms, for the Hellinger and $L^1$ distances on probability density functions on $[0; 1)$ for arbitrary Hölder regularity $\alpha > 0$. This improves upon previous results for constructions related to the Pólya tree prior, whose optimality was only proven when $\alpha \leq 1$. Also, by adding a hyperprior on the trees’ depth, we obtain an adaptive version of the prior that does not require $\alpha$ to be specified to attain optimality.

Keywords: Bayesian nonparametrics; ensemble; forest; Pólya Tree

References


Asymptotically equivalent prediction in multivariate geostatistics

FRANÇOIS BACHOC¹,a, EMILIO PORCU²,b, MORENO BEVILACQUA³,c, REINHARD FURRER⁴,d and TARIK FAOUZI⁵,e

¹Institut de mathématique, UMR5219; Université de Toulouse; CNRS, UPS IMT, F-31062 Toulouse Cedex 9, France. afrancois.bachoc@math.univ-toulouse.fr
²Department of Mathematics, Khalifa University, Abu Dhabi, The United Arab Emirates; School of Computer Science and Statistics, Trinity College, Dublin. bporcue@tcd.ie
³Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Viña del Mar, Chile; ANID-Millennium Science Initiative Program-Millennium Nucleus Center for the Discovery of Structures in Complex Data. cmoreno.bevilacqua@uai.cl
⁴Department of Mathematics and Department of Computational Science, University of Zurich, 8057, Zurich, Switzerland. dreinhard.furrer@math.uzh.ch
⁵Department of Statistics, Universidad del Biobio, 4081112, Concepción, Chile. etfaouzi@ubiobio.cl

Cokriging is the common method of spatial interpolation (best linear unbiased prediction) in multivariate geostatistics. While best linear prediction has been well understood in univariate spatial statistics, the literature for the multivariate case has been elusive so far. The new challenges provided by modern spatial datasets, being typically multivariate, call for a deeper study of cokriging. In particular, we deal with the problem of misspecified cokriging prediction within the framework of fixed domain asymptotics. Specifically, we provide conditions for equivalence of measures associated with multivariate Gaussian random fields, with index set in a compact set of a \( d \)-dimensional Euclidean space. Such conditions have been elusive for over about 50 years of spatial statistics.

We then focus on the multivariate Matérn and Generalized Wendland classes of matrix valued covariance functions, that have been very popular for having parameters that are crucial to spatial interpolation, and that control the mean square differentiability of the associated Gaussian process. We provide sufficient conditions, for equivalence of Gaussian measures, relying on the covariance parameters of these two classes. This enables to identify the parameters that are crucial to asymptotically equivalent interpolation in multivariate geostatistics. Our findings are then illustrated through simulation studies.

Keywords: Cokriging; equivalence of Gaussian measures; fixed domain asymptotics; functional analysis; Generalized Wendland; Matérn; spectral analysis

References


Phase transitions in non-linear urns with interacting types

MARCELO COSTA\textsuperscript{1,a} and JONATHAN JORDAN\textsuperscript{2,b}

\textsuperscript{1}Universidad de Buenos Aires, Buenos Aires, Argentina. \textsuperscript{a}mrocha@dm.uba.ar
\textsuperscript{2}School of Mathematics and Statistics, University of Sheffield, Sheffield, U.K. \textsuperscript{b}jonathan.jordan@sheffield.ac.uk

We investigate reinforced non-linear urns with interacting types, and show that where there are three interacting types there are phenomena which do not occur with two types. In a model with three types where the interactions between the types are symmetric, we show the existence of a double phase transition with three phases: as well as a phase with an almost sure limit where each of the three colours is equally represented and a phase with almost sure convergence to an asymmetric limit, which both occur with two types, there is also an intermediate phase where both symmetric and asymmetric limits are possible. In a model with anti-symmetric interactions between the types, we show the existence of a phase where the proportions of the three colours cycle and do not converge to a limit, alongside a phase where the proportions of the three colours can converge to a limit where each of the three is equally represented.

Keywords: Non-linear urn; phase transition

References

Continuous-time digital search tree and a border aggregation model

SVANTE JANSON\textsuperscript{1,a} and DEBLEENA THACKER\textsuperscript{2,b}

\textsuperscript{1}Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden.\textsuperscript{a}svante.janson@math.uu.se

\textsuperscript{2}Department of Mathematics, NYU, Shanghai, 1555 Century Avenue, Pudong New District, Shanghai, China 200122.\textsuperscript{b}thackerdebleena@gmail.com

We consider the continuous-time version of the random digital search tree, and construct a coupling with a border aggregation model as studied in Thacker and Volkov (Ann. Appl. Probab. 28 (2018) 1604–1633), showing a relation between the height of the tree and the time required for aggregation. This relation carries over to the corresponding discrete-time models. As a consequence we find a very precise asymptotic result for the time to aggregation, using recent results by Drmota et al. (Random Structures Algorithms 58 (2021) 430–467) for the digital search tree.

Keywords: Random walk; digital search trees; border aggregation model

References

Matrix means and a novel high-dimensional shrinkage phenomenon

ASAD LODHIA\textsuperscript{1,a}, KEITH LEVIN\textsuperscript{2,b} and ELIZAVETA LEVINA\textsuperscript{3,c}

\textsuperscript{1}Broad Institute of MIT and Harvard, 415 Main Street, Cambridge, MA 02142-1027. \textsuperscript{a}asad.i.lodhia@gmail.com
\textsuperscript{2}Department of Statistics, University of Wisconsin-Madison, 1220 Medical Sciences Center, 1300 University Ave, Madison, WI 53706-1510. \textsuperscript{b}kdlevin@wisc.edu
\textsuperscript{3}Department of Statistics, University of Michigan, 323 West Hall, 1085 S. University Ave, Ann Arbor, MI 48109-1107. \textsuperscript{c}elevina@umich.edu

Many statistical settings call for estimating a population parameter, most typically the population mean, based on a sample of matrices. The most natural estimate of the population mean is the arithmetic mean, but there are many other matrix means that may behave differently, especially in high dimensions. Here we consider the matrix harmonic mean as an alternative to the arithmetic matrix mean. We show that in certain high-dimensional regimes, the harmonic mean yields an improvement over the arithmetic mean in estimation error as measured by the operator norm. Counter-intuitively, studying the asymptotic behavior of these two matrix means in a spiked covariance estimation problem, we find that this improvement in operator norm error does not imply better recovery of the leading eigenvector. We also show that a Rao-Blackwellized version of the harmonic mean is equivalent to a linear shrinkage estimator studied previously in the high-dimensional covariance estimation literature, while applying a similar Rao-Blackwellization to regularized sample covariance matrices yields a novel nonlinear shrinkage estimator. Simulations complement the theoretical results, illustrating the conditions under which the harmonic matrix mean yields an empirically better estimate.

\textit{Keywords:} Matrix means; positive definite cone; covariance estimation; shrinkage Estimators

\textbf{References}


A pseudo-marginal sequential Monte Carlo online smoothing algorithm

PIERRE GLOAGUEN\textsuperscript{1,a}, SYLVAIN LE CORFF\textsuperscript{2,b} and JIMMY OLSSON\textsuperscript{3,c}

\textsuperscript{1}AgroParisTech, Paris, France. \textsuperscript{a}pierre.gloaguem@agroparistech.fr
\textsuperscript{2}Télécom SudParis, Paris, France. \textsuperscript{b}sylvain_lecorff@telecom-sudparis.eu
\textsuperscript{3}KTH Royal Institute of Technology, Stockholm, Sweden. \textsuperscript{c}jimmyol@kth.se

We consider online computation of expectations of additive state functionals under general path probability measures proportional to products of unnormalised transition densities. These transition densities are assumed to be intractable but possible to estimate, with or without bias. Using pseudo-marginalisation techniques we are able to extend the particle-based, rapid incremental smoother (PaRIS) algorithm proposed in [Bernoulli 23(3) (2017) 1951–1996] to this setting. The resulting algorithm, which has a linear complexity in the number of particles and constant memory requirements, applies to a wide range of challenging path-space Monte Carlo problems, including smoothing in partially observed diffusion processes and models with intractable likelihood. The algorithm is furnished with several theoretical results, including a central limit theorem, establishing its convergence and numerical stability. Moreover, under strong mixing assumptions we establish a novel $O(n\epsilon)$ bound on the asymptotic bias of the algorithm, where $n$ is the path length and $\epsilon$ controls the bias of the transition-density estimators.

Keywords: Central limit theorem; exponential concentration; partially observed diffusions; particle smoothing; pseudo-marginal methods; sequential Monte Carlo methods

References

A pseudo-marginal SMC online smoothing algorithm


Deviation inequalities and Cramér-type moderate deviations for the explosive autoregressive process

HUI JIANG¹,a, YILONG WAN¹,b and GUANGYU YANG²,c

¹Department of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing, China.
²School of Mathematics and Statistics, Zhengzhou University, Zhengzhou, China.

This paper concerns the asymptotic properties of the quadratic functionals and associated ordinary least squares estimator in the explosive first-order Gaussian autoregressive process. By the deviation inequalities for multiple Wiener-Itô integrals and asymptotic analysis techniques, Cramér-type moderate deviations are achieved under the explosive and mildly explosive frameworks. As applications, the global and local powers for the unit root test are shown to approach one at exponential rates. Simulation experiments are conducted to confirm the theoretical results.

Keywords: Cramér-type moderate deviations; deviation inequality; explosive autoregressive process; multiple Wiener-Itô integrals; unit root

References

Deviation inequalities and CMDP for explosive AR process


Estimation of Wasserstein distances in the Spiked Transport Model

JONATHAN NILES-WEED\textsuperscript{1,a} and PHILIPPE RIGOLLET\textsuperscript{2,b}

\textsuperscript{1}Courant Institute of Mathematical Sciences & Center for Data Science, New York University, 251 Mercer Street, New York, NY 10012-1185, USA. \textsuperscript{2}Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA.

\textsuperscript{a}jnw@cims.nyu.edu \textsuperscript{b}rigollet@math.mit.edu

Abstract. We propose a new statistical model, the \textit{spiked transport model}, which formalizes the assumption that two probability distributions differ only on a low-dimensional subspace. We study the minimax rate of estimation for the Wasserstein distance under this model and show that this low-dimensional structure can be exploited to avoid the curse of dimensionality. As a byproduct of our minimax analysis, we establish a lower bound showing that, in the absence of such structure, the plug-in estimator is nearly rate-optimal for estimating the Wasserstein distance in high dimension. We also give evidence for a statistical-computational gap and conjecture that any computationally efficient estimator is bound to suffer from the curse of dimensionality.

Keywords: Wasserstein distance; optimal transport; high-dimensional statistics

References


A presmoothing approach for estimation in the semiparametric Cox mixture cure model

ENI MUSTA 1,2,b, VALENTIN PATILEA 3,c and INGRID VAN KEILEGOM 1,a

1 ORSTAT, KU Leuven, Belgium. 2 Korteweg de Vries Institute for Mathematics, University of Amsterdam, Netherlands. 3 CREST, Ensai, France.

A challenge when dealing with survival analysis data is accounting for a cure fraction, meaning that some subjects will never experience the event of interest. Mixture cure models have been frequently used to estimate both the probability of being cured and the time to event for the susceptible subjects, by usually assuming a parametric (logistic) form of the incidence. We propose a new estimation procedure for a parametric cure rate that relies on a preliminary smooth estimator and is independent of the model assumed for the latency. On a second stage one can assume a semiparametric model for the latency and estimate also the survival distribution of the uncured subject. For the particular case of the logistic/Cox model, we investigate the theoretical properties of the estimators and show through simulations that presmoothing leads to more accurate results compared to the maximum likelihood estimator. To illustrate the practical use, we apply the new estimation procedure to two studies of melanoma survival data.

Keywords: Cure models; kernel smoothing; logistic model; survival analysis

References


Asymptotic behavior of a low-temperature non-cascading 2-GREM dynamics at extreme time scales

LUIZ RENATO FONTES¹,a, SUSANA FRÓMETA²,c and LEONEL ZUAZNÁBAR¹,b

¹USP - Instituto de Matemática e Estatística, R. do Matão, 1010 - Butantã, São Paulo - SP, CEP: 05508-090, Brasil. ¹lrfontes@usp.br, ¹lzuaznabar@ime.usp.br
²UFRGS - Instituto de Matemática e Estatística, Campus do Vale, Av. Bento Gonçalves, Porto Alegre - RS, 9500, CEP: 91509-900, Brasil. ²susana.frometa@ufrgs.br

We derive the scaling limit for the Hierarchical Random Hopping dynamics for the non cascading 2-GREM at low temperatures and time scales where the dynamics is close to equilibrium. The fine tuning phenomenon plays a role (under certain choices of parameters of the model), yielding three dynamical regimes. In contrast to the cascading case, the pairs of first and second level energies have fluctuations which scale with the volume, and this leads to a family of time scales where we see the dynamics moving through only a part of the full low-temperature energy landscape.

Keywords: GREM; random hopping dynamics; low temperature; fine tuning temperature; scaling limit; extreme time-scale; ergodic time-scale; K process; spin glasses

References

Non-cascading 2-GREM dynamics at extreme time scales


Estimation of the $\ell_2$-norm and testing in sparse linear regression with unknown variance

ALEXANDRA CARPENTIER$^{1,a}$, OLIVIER COLLIER$^{2,b}$, LAETITIA COMMINGES$^{3,c}$, ALEXANDRE B. TSYBAKOV$^{4,d}$ and YUHAO WANG$^{5,e}$

$1$ Institut für Mathematik, Universität Potsdam, Potsdam, Germany. $^a$carpentier@uni-potsdam.de
$2$ Modal’X, Université Paris-Nanterre, Nanterre and CREST, Paris, France. $^b$olivier.collier@parisnanterre.fr
$3$ CEREMADE, Université Paris-Dauphine and CREST, Paris, France. $^c$laetitia.comminges@dauphine.psl.eu
$4$ CREST, ENSAE and Institut Polytechnique de Paris, Paris, France. $^d$alexandre.tsybakov@ensae.fr
$5$ Tsinghua University, Beijing, China and Shanghai Qi Zhi Institute, Shanghai, China. $^e$yuhaow@tsinghua.edu.cn

We consider the related problems of estimating the $\ell_2$-norm and the squared $\ell_2$-norm in sparse linear regression with unknown variance, as well as the problem of testing the hypothesis that the regression parameter is null under sparse alternatives with $\ell_2$ separation. We establish the minimax optimal rates of estimation (respectively, testing) in these three problems.

Keywords: Sparse linear regression; signal detection; non-linear functional estimation

References


\[ \ell_2 \text{-norm estimation and testing in sparse linear regression} \]


A unified performance analysis of likelihood-informed subspace methods

TIANGANG CUI\textsuperscript{1,a} and XIN T. TONG\textsuperscript{2,b}

\textsuperscript{1}Monash University, School of Mathematics. \textsuperscript{a}tiangang.cui@monash.edu
\textsuperscript{2}National University of Singapore, Department of Mathematics. \textsuperscript{b}mattxin@nus.edu.sg

The likelihood-informed subspace (LIS) method offers a viable route to reducing the dimensionality of high-dimensional probability distributions arising in Bayesian inference. LIS identifies an intrinsic low-dimensional linear subspace where the target distribution differs the most from some tractable reference distribution. Such a subspace can be identified using the leading eigenvectors of a Gram matrix of the gradient of the log-likelihood function. Then, the original high-dimensional target distribution is approximated through various forms of marginalization of the likelihood function, in which the approximated likelihood only has support on the intrinsic low-dimensional subspace. This approximation enables the design of inference algorithms that can scale sub-linearly with the apparent dimensionality of the problem. Intuitively, the accuracy of the approximation, and hence the performance of the inference algorithms, are influenced by three factors—the dimension truncation error in identifying the subspace, Monte Carlo error in estimating the Gram matrices, and Monte Carlo error in constructing marginalizations. This work establishes a unified framework to analyze each of these three factors and their interplay. Under mild technical assumptions, we establish error bounds for a range of existing dimension reduction techniques based on the principle of LIS. Our error bounds also provide useful insights into the accuracy of these methods. In addition, we analyze the integration of LIS with sampling methods such as Markov Chain Monte Carlo (MCMC) and sequential Monte Carlo (SMC). We also demonstrate the applicability of our analysis on a linear inverse problem with Gaussian prior, which shows that all the estimates can be dimension-independent if the prior covariance is a trace-class operator. Finally, we demonstrate various aspects of our theoretical claims on two nonlinear inverse problems.

Keywords: Dimension reduction; approximation error; likelihood informed subspace; Monte Carlo estimation

References


1350-7265 © 2022 ISI/BS


Let $\pi \in \Pi(\mu, \nu)$ be a coupling between two probability measures $\mu$ and $\nu$ on a Polish space. In this article we propose and study a class of nonparametric measures of association between $\mu$ and $\nu$, which we call Wasserstein correlation coefficients. These coefficients are based on the Wasserstein distance between $\nu$ and the disintegration $\pi_{x,1}$ of $\pi$ with respect to the first coordinate. We also establish basic statistical properties of this new class of measures: we develop a statistical theory for strongly consistent estimators and determine their convergence rate in the case of compactly supported measures $\mu$ and $\nu$. Throughout our analysis we make use of the so-called adapted/bicausal Wasserstein distance, in particular we rely on results established in [Backhoff, Bartl, Beiglböck, Wiesel. Estimating processes in adapted Wasserstein distance. 2020]. Our approach applies to probability laws on general Polish spaces.

**Keywords:** Independence; measure of association; correlation; optimal transport; empirical measure; (bicausal/adapted) Wasserstein distance

**References**


Measuring association with Wasserstein distances


Local scaling limits of Lévy driven fractional random fields

VYTAUTE PILIPAUSKAITĖ¹,a and DONATAS SURGAILIS²,b

¹University of Luxembourg, Department of Mathematics, 6 Avenue de la Fonte, 4364 Esch-sur-Alzette, Luxembourg. avytaute.pilipauskaite@gmail.com
²Vilnius University, Faculty of Mathematics and Informatics, Naugarduko 24, 03225 Vilnius, Lithuania. bdonatas.surgailis@mif.vu.lt

We obtain a complete description of local anisotropic scaling limits for a class of fractional random fields $X$ on $\mathbb{R}^2$ written as stochastic integral with respect to infinitely divisible random measure. The scaling procedure involves increments of $X$ over points the distance between which in the horizontal and vertical directions shrinks as $O(\lambda^\alpha)$ and $O(\lambda^{\alpha'})$ respectively as $\lambda \downarrow 0$, for some $\gamma > 0$. We consider two types of increments of $X$: usual increment and rectangular increment, leading to the respective concepts of $\gamma$-tangent and $\gamma$-rectangent random fields. We prove that for above $X$ both types of local scaling limits exist for any $\gamma > 0$ and undergo a transition, being independent of $\gamma > \gamma_0$ and $\gamma < \gamma_0$, for some $\gamma_0 > 0$; moreover, the ‘unbalanced’ scaling limits ($\gamma \neq \gamma_0$) are $(H_1,H_2)$-multi self-similar with one of $H_i$, $i = 1,2$, equal to 0 or 1. The paper extends Pilipauskaitė and Surgailis (Stochastic Process. Appl. 127 (2017) 2751–2779) and Surgailis (Stochastic Process. Appl. 130 (2020) 7518–7546) on large-scale anisotropic scaling of random fields on $\mathbb{Z}^2$ and Benassi et al. (Bernoulli 10 (2004) 357–373) on 1-tangent limits of isotropic fractional Lévy random fields.

Keywords: Fractional random field; local anisotropic scaling limit; rectangular increment; Lévy random measure; scaling transition; multi self-similar random field

References

Local scaling limits of Lévy driven fractional RFs


Robust importance sampling with adaptive winsorization

PAULO ORENSTEIN

Instituto de Matemática Pura e Aplicada, Estr. Dona Castorina, 110, Rio de Janeiro 22460-320, Brazil.
a pauloo@impa.br

Importance sampling is a widely used technique to estimate properties of a distribution. The resulting estimator is unbiased but may have huge, potentially infinite, variance. This paper proposes trading-off some bias for variance by adaptively winsorizing the importance sampling estimator. The novel procedure is based on the Balancing Principle (or Lepskii’s Method). As a consequence, it offers a principled way to perform winsorization with finite-sample guarantees in the form of an oracle inequality. In various examples, the proposed estimator is shown to have smaller mean squared error and mean absolute deviation than leading alternatives such as the traditional importance sampling estimator or winsorizing it via cross-validation.

Keywords: Importance sampling; Winsorization; Balancing Principle; Lepskii’s method; robustness

References


Upper functions for sample paths of Lévy(-type) processes

FRANZISKA KÜHN

TU Dresden, Fakultät Mathematik, Institut für Mathematische Stochastik, 01062 Dresden, Germany.

franziska.kuehn1@tu-dresden.de

We study the small-time asymptotics of sample paths of Lévy processes and Lévy-type processes. Namely, we investigate under which conditions the limit

$$\limsup_{t \to 0} \frac{1}{f(t)}|X_t - X_0|$$

is finite resp. infinite with probability 1. We establish integral criteria in terms of the infinitesimal characteristics and the symbol of the process. Our results apply to a wide class of processes, including solutions to Lévy-driven SDEs and stable-like processes. For the particular case of Lévy processes, we recover and extend earlier results from the literature. Moreover, we present a new maximal inequality for Lévy-type processes, which is of independent interest.

Keywords: Lévy process; Feller process; martingale problem; sample path behaviour; small-time asymptotics; upper function; maximal inequality

References


Upper functions for Lévy(-type) processes


Quantifying deviations from separability in space-time functional processes

HOLGER DETTE\textsuperscript{a}, GAUTHIER DIERICKX\textsuperscript{b} and TIM KUTTA\textsuperscript{c}

Ruhr-Universität Bochum, D-44780 Bochum. \textsuperscript{a}holger.dette@rub.de, \textsuperscript{b}gauthier.dierickx@rub.de, \textsuperscript{c}tim.kutta@rub.de

The estimation of covariance operators of spatio-temporal data is in many applications only computationally feasible under simplifying assumptions, such as separability of the covariance into strictly temporal and spatial factors. Powerful tests for this assumption have been proposed in the literature. However, as real world systems, such as climate data are notoriously inseparable, validating this assumption by statistical tests, seems inherently questionable. In this paper we present an alternative approach: By virtue of separability measures, we quantify how strongly the data’s covariance operator diverges from a separable approximation. Confidence intervals localize these measures with statistical guarantees. This method provides users with a flexible tool, to weigh the computational gains of a separable model against the associated increase in bias. As separable approximations we consider the established methods of partial traces and partial products, and develop weak convergence principles for the corresponding estimators. Moreover, we also prove such results for estimators of optimal, separable approximations, which are arguably of most interest in applications. In particular we present for the first time statistical inference for this object, which has been confined to estimation previously.

Besides confidence intervals, our results encompass tests for approximate separability. All methods proposed in this paper are free of nuisance parameters and do not require computationally expensive resampling procedures. A simulation study underlines the advantages of our approach and its applicability is demonstrated by the investigation of German annual temperature data.

Keywords: Space-time processes; approximate separability; partial traces; optimal approximation

References


Tracy-Widom limit for the largest eigenvalue of high-dimensional covariance matrices in elliptical distributions

WEN JUN\textsuperscript{a}, XIE JIAHUI\textsuperscript{b}, YU LONG\textsuperscript{c} and ZHOU WANG\textsuperscript{d}

Department of Statistics and Data Science, National University of Singapore, Singapore. \textsuperscript{a}jun.wen@u.nus.edu, \textsuperscript{b}jiahui.xie@u.nus.edu, \textsuperscript{c}stayl@nus.edu.sg, \textsuperscript{d}wangzhou@nus.edu.sg

Let $X$ be an $M \times N$ random matrix consisting of independent $M$-variate elliptically distributed column vectors $x_1, \ldots, x_N$ with general population covariance matrix $\Sigma$. In the literature, the quantity $XX^\ast$ is referred to as the sample covariance matrix after scaling, where $X^\ast$ is the transpose of $X$. In this article, we prove that the limiting behavior of the scaled largest eigenvalue of $XX^\ast$ is universal for a wide class of elliptical distributions, namely, the scaled largest eigenvalue converges weakly to the same limit regardless of the distributions that $x_1, \ldots, x_N$ follow as $M, N \to \infty$ with $M/N \to \phi_0 > 0$ if the weak fourth moment of the radius of $x_1$ exists. In particular, via comparing the Green function with that of the sample covariance matrix of multivariate normally distributed data, we conclude that the limiting distribution of the scaled largest eigenvalue is the celebrated Tracy-Widom law.

Keywords: Sample covariance matrices; elliptical distributions; edge universality; Tracy-Widom distribution; tail probability

References


Nonparametric inference for reversed mean models with panel count data

LI LIU¹,a, WEN SU²,b, GUOSHENG YIN²,c, XINGQIU ZHAO³,d and YING ZHANG⁴,e

¹School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei, 430072, China. a liu.math@whu.edu.cn
²Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong. b jenna.wen.su@connect.hku.hk, c gyin@hku.hk
³Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong. d xingqiu.zhao@polyu.edu.hk
⁴Department of Biostatistics, University of Nebraska Medical Center, Omaha, NE, USA. e ying.zhang@unmc.edu

Panel count data typically refer to data arising from studies with recurrent events, in which subjects are observed only at discrete time points rather than under continuous observations. We investigate a general situation where a recurrent event process is eventually truncated by an informative terminal event and we are particularly interested in behaviors of the recurrent event process near the terminal event. We propose a reversed mean model for estimating the mean function of the recurrent event process. We develop a two-stage sieve likelihood-based method to estimate the mean function, which overcomes the computational difficulties arising from a nuisance functional parameter involved in the likelihood. The consistency and the convergence rate of the two-stage estimator are established. Allowing for the convergence rate slower than the standard rate, we develop the general weak convergence theory of M-estimators with a nuisance functional parameter, and then apply it to the proposed estimator for deriving the asymptotic normality. Furthermore, a class of two-sample tests is developed. The proposed methods are evaluated with extensive simulation studies and illustrated with panel count data from the Chinese Longitudinal Healthy Longevity Study.

Keywords: Nonparametric tests; recurrent events; reversed mean model; terminal event

References


Coverage error optimal confidence intervals for local polynomial regression

SEBASTIAN CALONICO\textsuperscript{1,a}, MATIAS D. CATTANEO\textsuperscript{2,b} and MAX H. FARRELL\textsuperscript{3,c}

\textsuperscript{1}Department of Health Policy and Management, Columbia University, New York, New York, U.S.A. 
\textsuperscript{a}sebastian.calonico@columbia.edu
\textsuperscript{2}Department of Operations Research and Financial Engineering, Princeton University, Princeton, New Jersey, U.S.A. \textsuperscript{b}cattaneo@princeton.edu
\textsuperscript{3}Booth School of Business, University of Chicago, Chicago, Illinois, U.S.A. \textsuperscript{c}max.farrell@chicagobooth.edu

This paper studies higher-order inference properties of nonparametric local polynomial regression methods under random sampling. We prove Edgeworth expansions for $t$ statistics and coverage error expansions for interval estimators that (i) hold uniformly in the data generating process, (ii) allow for the uniform kernel, and (iii) cover estimation of derivatives of the regression function. The terms of the higher-order expansions, and their associated rates as a function of the sample size and bandwidth sequence, depend on the smoothness of the population regression function, the smoothness exploited by the inference procedure, and on whether the evaluation point is in the interior or on the boundary of the support. We prove that robust bias corrected confidence intervals have the fastest coverage error decay rates in all cases, and we use our results to deliver novel, inference-optimal bandwidth selectors. The main methodological results are implemented in companion \texttt{R} and \texttt{Stata} software packages.

{	extit{Keywords:} Edgeworth expansion; Cramér condition; nonparametric regression; robust bias correction; bandwidth selection; optimal inference; minimax bound}

\textbf{References}


1350-7265 © 2022 ISI/BS
Coverage error optimality


Graphical modeling of stochastic processes driven by correlated noise

SØREN WENGEL MOGENSEN\textsuperscript{1,2,3,a} and NIELS RICHARD HANSEN\textsuperscript{1,b}

\textsuperscript{1}Department of Mathematical Sciences, University of Copenhagen, Copenhagen, Denmark.
\textsuperscript{a}soren.wengel_mogensen@control.lth.se, \textsuperscript{b}niels.r.hansen@math.ku.dk
\textsuperscript{2}Section for Cognitive Systems, Technical University of Denmark, Kgs. Lyngby, Denmark
\textsuperscript{3}Current affiliation: Department of Automatic Control, Lund University, Lund, Sweden

We study a class of graphs that represent local independence structures in stochastic processes allowing for correlated noise processes. Several graphs may encode the same local independencies and we characterize such equivalence classes of graphs. In the worst case, the number of conditions in our characterizations grows superpolynomially as a function of the size of the node set in the graph. We show that deciding Markov equivalence of graphs from this class is coNP-complete which suggests that our characterizations cannot be improved upon substantially. We prove a global Markov property in the case of a multivariate Ornstein-Uhlenbeck process which is driven by correlated Brownian motions.

Keywords: Graphical models; stochastic processes; local independence; Markov equivalence; Ornstein–Uhlenbeck processes

References


[38] Mogensen, S.W. and Hansen, N.R. Supplement to “Graphical modeling of stochastic processes driven by correlated noise.” https://doi.org/10.3150/21-BEJ1446SUPP


Conditional regression for single-index models

ALESSANDRO LANTERI¹,a, MAURO MAGGIONI²,b and STEFANO VIGOGNA³,c

¹DEMM, Università degli Studi di Milano, Milano, Italy. ²Department of Mathematics and Department of Applied Mathematics & Statistics, Johns Hopkins University, Baltimore, USA. ³MaLGa, DIBRIS, Università degli Studi di Genova, Genova, Italy.

The single-index model is a statistical model for intrinsic regression where responses are assumed to depend on a single yet unknown linear combination of the predictors, allowing to express the regression function as $\mathbb{E}[Y|X] = f(\langle v, X \rangle)$ for some unknown index vector $v$ and link function $f$. Conditional methods provide a simple and effective approach to estimate $v$ by averaging moments of $X$ conditioned on $Y$, but depend on parameters whose optimal choice is unknown and do not provide generalization bounds on $f$. In this paper we propose a new conditional method converging at $\sqrt{n}$ rate under an explicit parameter characterization. Moreover, we prove that polynomial partitioning estimates achieve the 1-dimensional min-max rate for regression of Hölder functions when combined to any $\sqrt{n}$-convergent index estimator. Overall this yields an estimator for dimension reduction and regression of single-index models that attains statistical optimality in quasilinear time.

Keywords: Single-index model; dimension reduction; nonparametric regression; finite-sample bounds

References


Conditional regression for single-index models


