Springer Proceedings in Mathematics & Statistics

Jaromír Antoch
Jana Jurečková
Matúš Maciak
Michal Pešta *Editors*

Analytical Methods in Statistics

AMISTAT, Prague, November 2015

Springer

Springer Proceedings in Mathematics & Statistics

Volume 193

Springer Proceedings in Mathematics & Statistics

This book series features volumes composed of selected contributions from workshops and conferences in all areas of current research in mathematics and statistics, including operation research and optimization. In addition to an overall evaluation of the interest, scientific quality, and timeliness of each proposal at the hands of the publisher, individual contributions are all refereed to the high quality standards of leading journals in the field. Thus, this series provides the research community with well-edited, authoritative reports on developments in the most exciting areas of mathematical and statistical research today.

More information about this series at http://www.springer.com/series/10533

Jaromír Antoch · Jana Jurečková
Matúš Maciak · Michal Pešta
Editors

Analytical Methods in Statistics

AMISTAT, Prague, November 2015

Springer

Editors
Jaromír Antoch
Department of Probability and Mathematical Statistics
Charles University
Prague
Czech Republic

Matúš Maciak
Department of Probability and Mathematical Statistics
Charles University
Prague
Czech Republic

Jana Jurečková
Department of Probability and Mathematical Statistics
Charles University
Prague
Czech Republic

Michal Pešta
Department of Probability and Mathematical Statistics
Charles University
Prague
Czech Republic

ISSN 2194-1009 ISSN 2194-1017 (electronic)
Springer Proceedings in Mathematics & Statistics
ISBN 978-3-319-51312-6 ISBN 978-3-319-51313-3 (eBook)
DOI 10.1007/978-3-319-51313-3

Library of Congress Control Number: 2016960726

Mathematics Subject Classification (2010): 62E20, 62F05, 62F12, 62Jxx, 62G08, 60Gxx, 62Mxx, 62F40

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This proceedings volume grew out of the Workshop Analytical Methods in Statistics (AMISTAT 2015) which was held in Prague during November 10–13, 2015. The workshop was organized by the Fulbright Commission in Prague, which hosted Prof. Abram Kagan as the Fulbright Distinguished Chair at the Charles University, together with the American Center in Prague which made their lecture halls available for the workshop; our special thanks belong to both, the American Center and the Fulbright Commission.

The workshop brought together many interesting points and discussions on statistical decisions, arising from new problems everyday. We appreciate that young people could meet famous experts in the area and listen to their talks. The contributions dealing with current points of interest and with exciting open problems were presented by scholars from Belgium, France, Germany, India, Israel, Norway, Russia, Serbia, Slovakia, Sweden, UK, USA, and the Czech Republic. A part of these contributions, by those authors who considered their analyses as temporarily finished, is contained in the present book. The joint motto of the talks and all discussions was the "analytical statistics". It emphasizes that the statistics provides mathematicians with challenging and exciting problems, because it obtains its problems from the real-life activities. For such problems one can rarely determine any axioms, and new problems appear every day. The statisticians utilize knowledge from all parts of mathematics, from those very abstract to numerical computation and interpretation of the results. Moreover, every statistician is expected to find a solution to a real problem and would not afford to reply that the optimal solution does not exist. He/she should look at least for a solution optimal under acceptable constraints, and to find it is again a challenge. Even thinking mathematically he/she has a feedback of practicality of the conclusions in mind.

We deeply appreciate the help and effort of Dr. Veronika Rosteck, Springer Editor of Statistics, for her encouragement, help and for possibility of publication of this book. We thank all the authors of the chapters and all referees for their work and consideration.

Prague, Czech Republic
September 2016

Jaromír Antoch
Jana Jurečková
Matúš Maciak
Michal Pešta

Contents

Contributors

M. Rauf Ahmad Department of Statistics, Uppsala University, Uppsala, Sweden

Michel Broniatowski Laboratoire de Statistique Théorique et Appliquée, Université Pierre et Marie Curie, Paris, France

Zdeněk Hlávka Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic

Xiyu Jiao Department of Economics, University of Oxford and Mansfield College, Oxford, UK

Jana Jurečková Faculty of Mathematics and Physics, Department of Probability and Statistics, Charles University, Prague, Czech Republic

Lev B. Klebanov Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic

Hira L. Koul Department of Statistics and Probability, Michigan State University, East Lansing, MI, USA

Matúš Maciak Department of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic

Bent Nielsen Department of Economics, University of Oxford and Nuffield College, Oxford, UK

Michal Pešta Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic

Alexander Sakhanenko Sobolev Institute of Mathematics, Novosibirsk, Russia

Lihong Wang Department of Mathematics, Nanjing University, Nanjing, People's Republic of China

Silvelyn Zwanzig Department of Mathematics, Uppsala University, Uppsala, Sweden

A Weighted Bootstrap Procedure for Divergence Minimization Problems

Michel Broniatowski

Abstract Sanov-type results hold for some weighted versions of empirical measures, and the rates for those Large Deviation principles can be identified as divergences between measures, which in turn characterize the form of the weights. This correspondence is considered within the range of the Cressie–Read family of statistical divergences, which covers most of the usual statistical criterions. We propose a weighted bootstrap procedure in order to estimate these rates. To any such rate we produce an explicit procedure which defines the weights, therefore replacing a variational problem in the space of measures by a simple Monte Carlo procedure.

Keywords Divergence · Optimization · Bootstrap · Monte Carlo · Large deviation · Weighted empirical measure · Conditional Sanov theorem

1 The Scope of This Paper

Recall that a sequence of random elements X_n with values in a measurable space $(T, \mathscr{T})$ satisfies a Large Deviation Principle with rate Φ whenever, for all measurable set $\Omega \subset T$ it holds

$$\begin{aligned}\Phi\left(int\left(\Omega\right)\right) &\leq -\lim\inf_{n\to\infty}\frac{1}{n}\log P\left(X_n\in\Omega\right)\\ &\leq -\lim\sup_{n\to\infty}\frac{1}{n}\log P\left(X_n\in\Omega\right)\leq\Phi\left(cl\left(\Omega\right)\right)\end{aligned}$$

where $int\left(\Omega\right)$ (resp. $cl\left(\Omega\right)$) denotes the interior (resp. the closure) of Ω in T and $\Phi(\Omega) := \inf\left\{\Phi(t); t\in\Omega\right\}$. The σ-field $\mathscr{T}$ is the Borel one defined by a given basis on T. For subsets Ω in T such that

M. Broniatowski (✉)
Laboratoire de Statistique Théorique et Appliquée, Université Pierre et Marie Curie,
Boîte Courrier 158, 4 place Jussieu, 75252 Paris Cedex 05, France
e-mail: michel.broniatowski@courriel.upmc.fr

J. Antoch et al. (eds.), *Analytical Methods in Statistics*, Springer Proceedings
in Mathematics & Statistics 193, DOI 10.1007/978-3-319-51313-3_1

$$\Phi\left(int\left(\Omega\right)\right)=\Phi\left(cl\left(\Omega\right)\right) \tag{1}$$

it follows by inclusion that

$$\begin{aligned}-\lim_{n\to\infty}\frac{1}{n}\log P\left(X_n\in\Omega\right)&=\Phi\left(int\left(\Omega\right)\right) \\ &=\Phi\left(cl\left(\Omega\right)\right)=\inf_{t\in\Omega}\Phi(t)=\Phi(\Omega).\end{aligned} \tag{2}$$

Assume that we are given such a family of random elements $X_1, X_2, \ldots$ together with a set $\Omega\subset T$ which satisfies (1). Suppose that we are interested in estimating $\Phi\left(\Omega\right)$. Then, whenever we are able to simulate a family of replicates $X_{n,1},\ldots,X_{n,K}$ such that $P\left(X_n\in\Omega\right)$ can be approximated by the frequency of those $X_{n,i}$'s in Ω, say

$$f_{n,K}\left(\Omega\right):=\frac{1}{K}card\left(i:X_{n,i}\in\Omega\right) \tag{3}$$

a natural estimator of $\Phi\left(\Omega\right)$ writes

$$\Phi_{n,K}\left(\Omega\right):=-\frac{1}{n}\log f_{n,K}\left(\Omega\right). \tag{4}$$

The rationale for this proposal is that visits of Ω by the random elements $X_{n,j}$'s tend to concentrate on the most favorable domain in Ω, namely where Φ assumes its minimal value in Ω, since $(\exp -n\Phi(x))\,dx$ is a good first-order approximation for the probability that X_n belongs to a neighborhood of x with volume dx. We have substituted the approximation of the variational problem $\Phi\left(\Omega\right):=\inf\left(\Phi\left(\omega\right),\omega\in\Omega\right)$ by a much simpler one, namely a Monte Carlo one, defined by (3). Notice further that we do not need to identify the set of points ω in Ω which minimize Φ; indeed there may be no such points even. Condition (1) provides an easy way to get statement (2), which yields to our estimates (4). Sometimes we may obtain (2) bypassing (1).

This program can be realized whenever we can identify the sequence of random elements X_i's for which, given the criterion Φ and the set Ω, the limit statement (2) holds. The present paper explores this approach in the case when the X_i's are empirical measures of some kind, and $\Phi(\Omega)$ writes $\phi\left(\Omega,P\right)$ which is the infimum of a divergence between some reference probability measure P and a class of probability measures Ω. This technique may lead to inferential procedures: for example assuming that $\Omega=\{Q_\theta\in\mathscr{M}_1,\theta\in\Theta\}$ is a statistical model such that $d(Q_\theta,P)\geq\varepsilon$ for some given distance d and some $\varepsilon>0$ and all θ in Θ, then minimizing a proxy of $\phi\left(\Omega_\theta,P\right)$ as obtained in this paper over θ provides minimum distance estimators of P within Ω.

The present paper presents estimators of $\phi\left(\Omega,P\right)$, focusing on their construction. We denote (P) the problem of finding an estimator for

$$\phi\left(\Omega,P\right) \tag{5}$$

where Ω is defined according to the context. But for simple convergence result of the proposed estimators, we do not provide finite sample or asymptotic properties of the estimators, which is postponed to future work; as seen later the method which we propose holds for rather general sets Ω; henceforth specific limit results of the estimator depend on the peculiar nature of the problem. Also the definition of the estimator through (4) may be changed using a better estimator of $P(X_n \in \Omega)$ than $f_{n,K}(\Omega)$, the naive one, which may have poor statistical performances and which may require a long runtime for calculation, since $(X_n \in \Omega)$ is a rare event; Importance Sampling procedures should be used. This is also out of the scope of this paper.

1.1 Existing Solutions for Similar Problems

Minimizing a divergence between an empirical measure pertaining to a data set and a class of distributions is somehow synonymous as estimating the parent distribution of the data (although other methods exist); for example the maximum likelihood method amounts to minimize the likelihood (or modified Kullback–Leibler) divergence between P_n and a parametrized model. Inspired by the celebrated Empirical Likelihood approach, empirical divergence methods aim at finding solutions of the minimization of the divergence between P_n and all distributions in Ω which are supported by the data points; see [1]. Those may exist or not, yielding (or not yielding) to the estimation of the minimum value of the divergence. Besides the fact that Ω may consists in distributions which cannot have the data as supporting points, the resulting equations for the solution of the problem may be intractable. Also there may be an infinity of solutions for this problem. The case when Ω is defined by conditions on moments of some $L-$statistics is illuminating in this respect; indeed the direct approach fails, and leads to a new problem, defining divergences between quantile measures (see [2]). Instead, looking first for some estimator of the infimum value of the divergence leads to a well posed problem of finding the set of minimizers, an algorithmic problem for which a solution can be obtained along the lines of the present paper. Once obtained the minimal value of the divergence, minimizers may sometimes be obtained by dichotomous search; this depends on the context.

2 Divergences

Let $(\mathscr{X}, \mathscr{B})$ be a measurable space and P be a given reference probability measure (p.m.) on $(\mathscr{X}, \mathscr{B})$. The set $\mathscr{X}$ is assumed to be a Polish space. Denote $\mathscr{M}$ the real vector space of all signed finite measures on $(\mathscr{X}, \mathscr{B})$ and $\mathscr{M}(P)$ the vector subspace of all signed finite measures absolutely continuous (a.c) with respect to (w.r.t.) P. Denote also $\mathscr{M}_1$ the set of all p.m.'s on $(\mathscr{X}, \mathscr{B})$ and $\mathscr{M}_1(P)$ the subset of all p.m.'s a.c w.r.t. P. Let φ be a proper[1] closed[2] convex function from $]-\infty, +\infty[$ to $[0, +\infty]$ with $\varphi(1) = 0$ and such that its domain $\mathrm{dom}\varphi := \{x \in \mathbb{R}$ such that $\varphi(x) < \infty\}$ is an interval with endpoints $a_\varphi < 1 < b_\varphi$ (which may be finite or infinite).

For any signed finite measure Q in $\mathscr{M}(P)$, a classical definition for the ϕ-divergence between Q and P is defined by

$$\phi(Q, P) := \int_{\mathscr{X}} \varphi\left(\frac{dQ}{dP}(x)\right) dP(x). \tag{6}$$

When Q is not a.c. w.r.t. P, we set $\phi(Q, P) = +\infty$; see [3]. The first definition of ϕ-divergences between p.m.'s were introduced by I. Csiszar in [4] as "f-divergences". Csiszar's definition of ϕ-divergences between p.m.'s requires a common dominating σ-finite measure λ for Q and P. Note that the two definitions of $\phi-$divergences coincide on the set of all p.m.'s a.c w.r.t. P and dominated by λ. The ϕ-divergences between any signed finite measure Q and a p.m. P were introduced by [5], which proposes the following definition

$$\phi(Q, P) := \int \varphi(q)\, dP + b_{\varphi^*}\sigma_Q^+(\mathscr{X}) - a_{\varphi^*}\sigma_Q^-(\mathscr{X}), \tag{7}$$

where

$$a_{\varphi^*} = \lim_{y\to-\infty} \frac{\varphi(y)}{y}, \quad b_{\varphi^*} = \lim_{y\to+\infty} \frac{\varphi(y)}{y}. \tag{8}$$

and

$$Q = qP + \sigma_Q, \quad \sigma_Q = \sigma_Q^+ - \sigma_Q^-$$

is the Lebesgue decomposition of Q, and the Jordan decomposition of the singular part σ_Q, respectively. Definitions (6) and (7) coincide when Q is a.c. w.r.t. P or when $a_\varphi = -\infty$ or $b_\varphi = +\infty$. Since we will consider optimization of $Q \mapsto \phi(Q, P)$ on sets of signed finite measures a.c. w.r.t. P, it is more adequate for our sake to use the definition (7).

For all p.m. P, the mappings $Q \in \mathscr{M} \mapsto \phi(Q, P)$ are convex and take nonnegative values. When $Q = P$ then $\phi(Q, P) = 0$. Furthermore, if the function $x \mapsto \varphi(x)$ is strictly convex on a neighborhood of $x = 1$, then the following basic property holds

$$\phi(Q, P) = 0 \text{ if and only if } Q = P. \tag{9}$$

All these properties are presented in [4, 6–8] Chap. 1, for ϕ-divergences defined on the set of all p.m.'s $\mathscr{M}_1$. When the ϕ-divergences are defined on $\mathscr{M}$, then the same properties hold making use of definition (7); see also [9].

[1] We say a function is proper if its domain is non void.

[2] The closedness of φ means that if a_φ or b_φ are finite numbers then $\varphi(x)$ tends to $\varphi(a_\varphi)$ or $\varphi(b_\varphi)$ when $x \downarrow a_\varphi$ or $x \uparrow b_\varphi$, respectively.

When defined on $\mathscr{M}_1$, the Kullback–Leibler (KL), modified Kullback–Leibler (KL_m), χ^2, modified χ^2 (χ^2_m), Hellinger (H), and L_1 divergences are respectively associated to the convex functions $\varphi(x) = x\log x - x + 1$, $\varphi(x) = -\log x + x - 1$, $\varphi(x) = \frac{1}{2}(x-1)^2$, $\varphi(x) = \frac{1}{2}(x-1)^2/x$, $\varphi(x) = 2(\sqrt{x}-1)^2$ and $\varphi(x) = |x-1|$. All those divergences except the L_1 one, belong to the class of power divergences introduced in [10] (see also [8] Chap. 2). They are defined through the class of convex functions

$$x \in]0, +\infty[\mapsto \varphi_\gamma(x) := \frac{x^\gamma - \gamma x + \gamma - 1}{\gamma(\gamma - 1)} \tag{10}$$

if $\gamma \in \mathbb{R} \setminus \{0, 1\}$, $\varphi_0(x) := -\log x + x - 1$ and $\varphi_1(x) := x\log x - x + 1$. (For all $\gamma \in \mathbb{R}$, we define $\varphi_\gamma(0) := \lim_{x\downarrow 0} \varphi_\gamma(x)$). So, the $KL-$divergence is associated to φ_1, the KL_m to φ_0, the χ^2 to φ_2, the χ^2_m to φ_{-1} and the Hellinger distance to $\varphi_{1/2}$.

Those divergence functions defined in (10) are the Cressie–Read divergence functions; see [10].

The Kullback–Leibler divergence (KL-divergence) is sometimes called Boltzmann Shannon relative entropy. It appears in the domain of large deviations and it is frequently used for reconstruction of laws, and in particular in the classical moment problem (see e.g. [5] and the references therein). The modified Kullback–Leibler divergence (KL_m-divergence) is sometimes called Burg relative entropy. It is frequently used in Statistics and it leads to efficient methods in statistical estimation and tests problems; in fact, the celebrate "maximum likelihood" method can be seen as an optimization problem of the KL_m-divergence between the discrete or continuous parametric model and the empirical measure associated to the data; see [11, 12]. On the other hand, the recent "empirical likelihood" method can also be seen as an optimization problem of the KL_m-divergence between some set of measures satisfying some linear constraints and the empirical measure associated to the data; see [13] and the references therein, [1, 14]. The Hellinger divergence is also used in Statistics, it leads to robust statistical methods in parametric and semi-parametric models; see [1, 15–17].

The power divergences functions $Q \in \mathscr{M}_1 \mapsto \phi_\gamma(Q, P)$ can be defined on the whole vector space of signed finite measures $\mathscr{M}$ via the extension of the definition of the convex functions φ_γ: For all $\gamma \in \mathbb{R}$ such that the function $x \mapsto \varphi_\gamma(x)$ is not defined on $]-\infty, 0[$ or defined but not convex on whole $\mathbb{R}$, we extend its definition as follows:

$$x \in]-\infty, +\infty[\mapsto \begin{cases} \varphi_\gamma(x) & \text{if } x \in [0, +\infty[, \\ +\infty & \text{if } x \in]-\infty, 0[. \end{cases} \tag{11}$$

Note that for the χ^2-divergence for instance, $\varphi_2(x) := \frac{1}{2}(x-1)^2$ is defined and convex on whole $\mathbb{R}$. This extension of the domain of the divergence functions φ_γ to $]-\infty, +\infty[$ implies that (8) is well defined, with $a_{\varphi^*} = +\infty$.

The conjugate (or Fenchel–Legendre transform) of φ will be denoted φ^*,

$$t \in \mathbb{R} \mapsto \varphi^*(t) := \sup_{x\in\mathbb{R}} \{tx - \varphi(x)\},$$

and the endpoints of domφ^* (the domain of φ^*) are a_{φ^*} and b_{φ^*} with $a_{\varphi^*} \leq b_{\varphi^*}$. Note that φ^* is a proper closed convex function. In particular, $a_{\varphi^*} < 0 < b_{\varphi^*}$, $\varphi^*(0) = 0$. By the closedness of φ, the conjugate φ^{**} of φ^* coincides with φ, i.e.,

$$\varphi^{**}(t) := \sup_{x \in \mathbb{R}} \left\{ tx - \varphi^*(x) \right\} = \varphi(t), \quad \text{for all } t \in \mathbb{R}.$$

For proper convex functions defined on $\mathbb{R}$ (endowed with the usual topology), the lower semi-continuity[3] and the closedness properties are equivalent.

We say that φ (resp. φ^*) is differentiable if it is differentiable on $]a_\varphi, b_\varphi[$ (resp. $]a_{\varphi^*}, b_{\varphi^*}[$), the interior of its domain. We say also that φ (resp. φ^*) is strictly convex if it is strictly convex on $]a_\varphi, b_\varphi[$ (resp. $]a_{\varphi^*}, b_{\varphi^*}[$).

The strict convexity of φ is equivalent to the condition that its conjugate φ^* is essentially smooth, i.e., differentiable with

$$\begin{array}{l} \lim_{t \downarrow a_{\varphi^*}} \varphi^{*\prime}(t) = -\infty \ \text{ if } \ a_{\varphi^*} > -\infty, \\ \lim_{t \uparrow b_{\varphi^*}} \varphi^{*\prime}(t) = +\infty \ \text{ if } \ b_{\varphi^*} < +\infty. \end{array}$$

Conversely, φ is essentially smooth if and only if φ^* is strictly convex; see e.g. [18] Sect. 26 for the proofs of these properties.

If φ is differentiable, we denote φ' the derivative function of φ, and we define $\varphi'(a_\varphi)$ and $\varphi'(b_\varphi)$ to be the limits (which may be finite or infinite) $\lim_{x \downarrow a_\varphi} \varphi'(x)$ and $\lim_{x \uparrow b_\varphi} \varphi'(x)$, respectively. We denote Imφ' the set of all values of the function φ', i.e., Im$\varphi' := \left\{ \varphi'(x) \text{ such that } x \in [a_\varphi, b_\varphi] \right\}$. If additionally the function φ is strictly convex, then φ' is increasing on $[a_\varphi, b_\varphi]$. Hence, it is a one-to-one function from $[a_\varphi, b_\varphi]$ onto Imφ'; we denote in this case φ'^{-1} the inverse function of φ' which is defined from Imφ' onto $[a_\varphi, b_\varphi]$.

Note that if φ is differentiable, then for all $x \in]a_\varphi, b_\varphi[$,

$$\varphi^* \left(\varphi'(x) \right) = x\varphi'(x) - \varphi(x). \tag{12}$$

If additionally φ is strictly convex, then for all $t \in$ Imφ' we have

$$\varphi^*(t) = t\varphi'^{-1}(t) - \varphi\left(\varphi'^{-1}(t) \right) \quad \text{and} \quad \varphi^{*\prime}(t) = \varphi'^{-1}(t).$$

On the other hand, if φ is essentially smooth, then the interior of the domain of φ^* coincides with that of Imφ', i.e., $\left(a_{\varphi^*}, b_{\varphi^*} \right) = \left(\varphi'(a_\varphi), \varphi'(b_\varphi) \right)$.

The domain of the ϕ-divergence will be denoted domϕ, i.e.,

$$\text{dom}\phi := \{ Q \in \mathscr{M} \text{ such that } \phi(Q, P) < \infty \}.$$

[3] We say a function φ is lower semi-continuous if the level sets $\{x \text{ such that } \varphi(x) \leq \alpha\}$, $\alpha \in \mathbb{R}$ are closed.

In the present paper we will deal with essentially smooth divergence functions, so that all the above properties are fulfilled.

3 Large Deviations for the Bootstrapped Empirical Measure

The present section aims at providing a solution to Problem (P) when Ω is a subset of $\mathscr{M}_1$, the class of all probability measures on $(\mathscr{X}, \mathscr{B})$. Such a goal will be achieved in two cases of interest, namely the Kullback–Leibler and the Likelihood divergence.

We first push forward some definition.

Let $Y, Y_1, Y_2, \ldots$ denote a sequence of nonnegative independent real- valued random variables with expectation 1. We assume that Y satisfies the so-called Cramer condition, namely that the set

$$\mathscr{N} := \left\{ t \in \mathbb{R} \text{ such that } \Lambda(t) := \log E e^{tY} < \infty \right\}$$

contains a neighborhood of 0 with non-void interior. By its very definition, $\mathscr{N}$ is an interval, say $\mathscr{N} := (a, b)$ which we assume to be open. We also assume that the strictly convex function Λ is a steep function, namely that $\lim_{t\to a} \Lambda(t) = \lim_{t\to b} \Lambda(t) = +\infty$. It will also be assumed that $t \to \Lambda'(t)$ parametrizes the convex hull of the support of the distribution of Y. We refer to [19] for those notions and conditions.

Consider now the weights W_i^n, $1 \le i \le n$ defined through

$$W_i^n := \frac{Y_i}{(1/n)\sum_{i=1}^n Y_i}$$

which define a vector of exchangeable variables $\left(W_1^n, \ldots, W_n^n\right)$ for all $n \ge 1$.

Define further the Legendre transform of Λ, say Λ^* which is a strictly convex function defined on Im Λ' by

$$\Lambda^*(x) := \sup_t tx - \Lambda(t).$$

We assume that we are given an array of observations $\left(x_i^n\right)_{i=1,\ldots,n,n\ge 1}$ in $\mathscr{X}$ which we assume to be "fair," meaning that there exists a probability measure P defined on $(\mathscr{X}, \mathscr{B})$ such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \delta_{x_i^n} = P. \tag{13}$$

When the observations are sampled under P we assume that the above condition (13) holds almost surely. We define the bootstrapped empirical measure of $\left(x_1^n, \ldots, x_n^n\right)$

by

$$P_n^W := \frac{1}{n}\sum_{i=1}^{n} W_i^n \delta_{x_i^n}.$$

Note that P_n^W is random due to the weights $W_1^n, \ldots, W_n^n$ and that the data set $x_1^n, \ldots, x_n^n$ is considered as nonrandom. The following result provides a Sanov-type LDP statement conditionally upon the array $\left(x_i^n\right)$ $1 \le i \le n, n \ge 1$. Assuming that Y has no atom at 0 and that $t \to \Lambda_Y(t)$ is steep at point

$$t^+ := \sup\{t : \Lambda_Y(t) < +\infty\}$$

with $t^+ > 0$, it holds

Theorem 1 *Under the above hypotheses and notation the sequence P_n^W obeys a LDP on the space of all probability measures on X equipped with the weak convergence topology with good rate function*

$$\phi(Q, P) := \begin{array}{ll} inf_{m>0} \int \Lambda^*\left(m\frac{dQ}{dP}(x)\right) dP(x) & ifQ << P \\ +\infty & otherwise \end{array} \tag{14}$$

Remark 1 This Theorem is a variation on Corollary 3.3 in [20]. Indeed it holds

$$\lim_{x\to-\infty} \Lambda_Y'(t) = \lim_{x\to-\infty} \left(\left(\Lambda^*\right)'\right)^{-1}(x) = 0$$

and

$$\lim_{x\to+\infty} \Lambda_Y'(t) = \lim_{x\to+\infty} \left(\left(\Lambda^*\right)'\right)^{-1}(x) = +\infty$$

The above Theorem does not meet our requirement that the rate should be a divergence between probability measures. Two cases of upmost interest however fulfill our quest.

We make use of independent copies of P_n^W, obtained as follows: consider

$$\left(Y_{1,1}, \ldots, Y_{1,n}\right), \ldots, \left(Y_{,1}, \ldots, Y_{K,n}\right)$$

where all the $Y_{i,j}$ are i.i.d. copies of Y, and

$$W_i^k := \frac{Y_{k,i}}{\sum_{i=1}^{k} Y_{k,i}},$$

$$P_{k,n}^W := \sum_{i=1}^{n} W_i^k \delta_{x_i^n}.$$

and for any set Ω in $\mathscr{M}_1$ define

$$P_{n,K}(\Omega) := \frac{1}{K} card\left(k \in \{1, \ldots, K\} : P_{k,n}^W \in \Omega\right) \tag{15}$$

and denote

$$L_{n,K}(\Omega) := -\frac{1}{n} \log P_{n,K}(\Omega). \tag{16}$$

3.1 Minimizing the Kullback–Leibler Divergence

Assume that the random variable Y is Poisson distributed with mean 1. Then

$$\Lambda^*(x) = x \log x - x + 1$$

which is the Kullback–Leibler divergence function. For any couple of probability measures (Q, P) it readily follows that the infimum upon m in (14) is reached at $m = \exp -KL(Q, P)$, which yields

$$inf_{m>0} \int \Lambda^* \left(m \frac{dQ}{dP}(x) \right) dP(x) = 1 - \exp -KL(Q, P). \tag{17}$$

It follows that the rate (14) takes the form

$$\phi(Q, P) = 1 - \exp -KL(Q, P)$$

and that

$$\phi(\Omega, P) = 1 - \exp -KL(\Omega, P)$$

Proposition 1 *Consider any set Ω of probability measures which satisfies*

$$KL(int\Omega, P) = KL(cl\Omega, P),$$

where $\mathscr{M}_1$ is endowed with the weak topology. Consider Y a r.v. with Poisson distribution with mean 1. Then the following expression

$$\widehat{KL}(\Omega, P) := -\log\left[1 - L_{n,K}(\Omega)\right]$$

estimates $KL(\Omega, P)$.

3.2 Minimizing the Likelihood Divergence

Let the r.v. Y have an exponential distribution with mean 1. Then

$$\Lambda^*(x) = -\log x + x - 1$$

which is the divergence function which defines the modified Kullback–Leibler divergence, also named as Likelihood divergence, since its minimization in statistically relevant contexts yields the celebrated maximum likelihood divergence estimators.

For all P and Q in $\mathscr{M}_1$ such that $KL_m(Q, P)$ is finite, the function $(0, 1) \ni m \to \int \varphi\left(m\frac{dQ}{dP}(x)\right) dP(x)$ is decreasing. Therefore the (14) takes the form

$$\phi\,(Q, P) = KL_m(Q, P)$$

and

$$\phi\,(\Omega, P) = KL_m(\Omega, P)$$

This yields an analogue of Proposition 1, namely

Proposition 2 *With the same notation and hypotheses as in Proposition 1, with Y a random variable with Exponential (1) distribution, the following expression*

$$\widehat{KL_m}\,(\Omega, P) := L_{n,K}\,(\Omega)$$

estimates $KL_m\,(\Omega, P)$.

Remark 2 When Y is exponentially distributed with expectation 1 then by Pyke's Theorem, the vector $\left(W_1^n, \ldots, W_n^n\right)$ coincides in distribution with

$$\left(U_{1,n}, U_{2,n} - U_{1,n}, \ldots, U_{n,n} - U_{n-1,n}\right),$$

the spacings of the ordered statistics $\left(U_{1,n}, U_{2,n}, \ldots, U_{n,n}\right)$ of n i.i.d. uniformly distributed r.v's on $(0, 1)$, with uniform distribution. This is indeed the simplest weighted bootstrap variation of P_n based on exchangeable weights.

4 Wild Bootstrap

We now consider other random elements whose visits in Ω will define estimators of minimum divergence between P and Ω for other useful divergence function, as the Chi-square, the Hellinger, etc.

We may consider some wild bootstrap versions, defining the wild empirical measure by

$$P_n^{Wild} := \frac{1}{n}\sum_{i=1}^{n} Y_i \delta_{x_{i,n}}$$

where the r.v's $Y_1, Y_2, \ldots$ are i.i.d. with common expectation 1. The use of the word "wild" is relevant: P_n^{Wild} is not merely a probability measure; it can even put negative masses on some points of its support, since the r.'s Y_i may assume negative values. We will be able to solve Problem (P) when Ω is a subset of $\mathscr{M}$, the class of all signed finite measures on $(\mathscr{X}, \mathscr{B})$. Thus the estimator of $\phi(\Omega, P)$ is typically smaller than the estimator of $\phi(\Omega \cap \mathscr{M}_1, P)$, which cannot be estimated using the results of this section, in contrast with just obtained in the previous section. Also we will need $\mathscr{X}$ to be a compact set.

We assume that the Cramer condition holds for Y and define, as above,

$$\Lambda_Y(t) := \log E \exp tY.$$

4.1 A Conditional LDP for the Wild Bootstrapped Empirical Measure

In this case we make use of the following result (see [21]) which holds when $\mathscr{X}$ is compact.

Theorem 2 *The wild empirical measure P_n^{Wild} obeys a LDP in the class of all signed finite measures endowed by the weak topology with good rate function $\phi(Q, P)$ defined in (7), where the function φ is defined by*

$$\varphi(x) := \Lambda^*(x) = \sup_t tx - \Lambda_Y(t).$$

Remark 3 Making use of the results in [21], we may consider the constant a_{φ^*} and b_{φ^*} in (7); by convexity, $\varphi^*(x) := \Lambda_Y(x)$. The LDP rate (7) writes

$$\phi(Q, P) := \int_{\mathscr{X}} \Lambda^* \left(\frac{dQ_a}{dP} \right) dP + \int_{\mathscr{X}} \rho \left(\frac{dQ_s}{d\theta} \right) d\theta$$

where

$$\rho(z) := \sup \{\lambda z : \lambda \in Dom \Lambda_Y\}$$

and θ is any real-valued nonnegative measure with respect to which Q_s is absolutely continuous. Choosing

$$\theta = \left| Q_s^+ - Q_s^- \right|$$

yields

$$\phi(Q, P) := \int_{\mathscr{X}} \Lambda^* \left(\frac{dQ_a}{dP} \right) dP + \rho(-1) Q_s^- (\mathscr{X}) + \rho(+1) Q_s^+ (\mathscr{X})$$

so that $a_{\varphi^*} = \inf \{t : \Lambda_Y(t) < \infty\}$ and $b_{\varphi^*} = \sup \{t : \Lambda_Y(t) < \infty\}$.

Remark 4 Theorem 2 has been proved by numerous authors, under various regularity conditions; see e.g. [21–23]. A strong result is as follows:

When Ω is a subset in $\mathscr{M}$ such that $\phi\left(cl\left(\Omega\right), P\right) = \phi\left(int\left(\Omega\right), P\right)$ holds in the $\tau-$topology, then

$$\lim_{n\to\infty} -\frac{1}{n} \log P\left(P_n^{Wild} \in \Omega\right) = \phi\left(\Omega, P\right). \tag{18}$$

However that $\tau-$open (resp. $\tau-$closed sets) are not necessarily weakly open (resp weakly closed); thus this latest result (18) is merely useful when Ω is defined as the pre-image of some open (closed) set by some $\tau-$continuous mapping from $(\mathscr{X}, \mathscr{B})$ onto some topological space; see Sect. 6.

4.2 Cressie–Read Divergences and Exponential Families

In this section we consider a reciprocal statement to Theorem 2. We first prove that any Cressie–Read divergence function as defined in (11) is the Fenchel–Legendre transform of some cumulant generating function Λ_Y for some r.v. Y. Henceforth, we state a one-to-one correspondence between the class of Cressie–Read divergence functions and the distribution of some Y which can be used in order to build a bootstrap empirical measure of the form P_n^{Wild}.

4.3 Natural Exponential Families and Their Variance Functions

We turn to some results due to Letac and Mora; see [24].

For μ a positive $\sigma-$finite measure on $\mathbb{R}$ define $\phi_\mu(t) := \int e^{tx} d\mu(x)$ and its domain $\mathscr{D}_\mu$, the set of all values of t such that $\phi_\mu(t)$ is finite, which is a convex (possibly void) subset of $\mathbb{R}$. Denote $k_\mu(t) := \log \phi_\mu(t)$ and let $m_\mu(t) := (d/dt)\, k_\mu(t)$ and $s_\mu^2(t) := \left(d^2/dt^2\right) k_\mu(t)$. Associated with μ is the Natural Exponential Family NEF(μ) of distributions

$$dP_t^\mu(x) := \frac{e^{tx} d\mu(x)}{\phi_\mu(t)}$$

which is indexed by t. It is a known fact that, denoting X_t a r.v. with distribution P_t^μ it holds $EX_t = m_\mu(t)$ and $VarX_t = s_\mu^2(t)$.The mapping $t \to m_\mu(t) := EX_t$ is a strictly increasing homeomorphism from $\mathbb{R}^+$ onto $\mathbb{R}^+$, with inverse $m_\mu^{\leftarrow}$.

The NEF(μ) is said to be *generated* by μ. The NEF(ν) generated by ν defined through

$$d\nu(x) = \exp(ax + b) d\mu(x) \tag{19}$$

coincides with NEF(μ), which yields to the definition of the NEF generated by the class of positive measures ν satisfying (19) for some constants a and b. Following [24] for the notation and main results the class of such measures will be denoted $\mathscr{B}$ and be called a *base* for NEF(μ), hence denoted NEF($\mathscr{B}$). Also it can be checked that the range of m_ν does not depend on the very choice of ν in $\mathscr{B}$, although its domain depends on ν. The range Im$m_{\mathscr{B}}$ of m_ν, which is the same for all ν in $\mathscr{B}$, is called the mean range of $\mathscr{B}$ since it depends only on the class of generating measures $\mathscr{B}$.

Defined on Im$m_{\mathscr{B}}$, the function

$$x \to V(x) := s_\mu^2 o m_\mu^{\leftarrow}(x)$$

is independent of the peculiar choice of μ in $\mathscr{B}$ (see [24]) and is therefore called the variance function of the NEF($\mathscr{B}$). It can be proved that the variance function characterizes the NEF. From the statistician point of view the functional form of the function V is of relevant interest: it corresponds to models for which regression of the variance on the mean is considered, which is a common feature in heteroscedastic models; see the seminal paper [25] which is at the origin of models characterized by V, and [26].

Starting with [27], a wide effort has been developed in order to characterize the basis of a NEF with given variance function.

4.4 Power Variance Functions and the Corresponding Natural Exponential Families

Power variance functions have been explored by various authors; see e.g. [24, 28], etc. Summarizing it holds (see [28]) the NEF with variance function $V(x) = Cx^\alpha$; for sake of brevity with respect to the sequel we denote $\alpha = 2 - \gamma$. NEF with variance function V are obtained through integration and identification of the resulting moment-generating function. They are generated as follows:

- For $\gamma < 0$ by stable distributions on $\mathbb{R}^+$ with characteristic exponent in $(0, 1)$. The resulting distributions define the Tweedie scale family, which we briefly describe in the next paragraph.
- For $\gamma = 0$ by the exponential distribution
- For $0 < \gamma < 1$ by Compound Gamma–Poisson distributions
- For $\gamma = 1$ by the Poisson distribution
- For $\gamma = 2$ by the normal distribution

Other values of γ do not yield NEF's.

4.4.1 The Tweedie Scale

Let Z be a r.v. with stable distribution on $\mathbb{R}^+$ with exponent τ, $0 < \tau < 1$. Denote p its density and $f(t) = E \exp itZ$ its characteristic function, which satisfies

$$f(t) = \exp\left\{iat - c\,|t|^{\tau}\left(1 + i\beta sign\,(t)\,\omega\,(t, \tau)\right)\right\}$$

where $a \in \mathbb{R}$, $c > 0$ and $\omega\,(t, \tau) = \tan\left(\frac{\pi\tau}{2}\right)$.

We consider the case when $\beta = 1$. It then holds:

For $Z_1, \ldots, Z_n$ n i.i.d. copies of Z,

$$\frac{Z_1 + \ldots + Z_n}{n^{1/\tau}} =_d Z$$

where the equality holds in distribution. The Laplace transform of p satisfies

$$\varphi(t) := \int_0^{\infty} e^{-tx}p(x)dx = e^{-t^{\tau}}$$

for all nonnegative value of t; see [29].

Associated with p is the Natural Exponential family (NEF) with basis p namely the densities defined for nonnegative t through

$$p_t(x) := e^{-tx}p(x)/e^{-t^{\tau}}$$

with support $\mathbb{R}^+$. For positive t, a r.v. X_t with density p_t has a moment-generating function $E \exp \lambda X_t$ which is finite in a non-void neighborhood of 0 and therefore has moments of any order.

Consider the density $p_1(x) = e^{-x+1}p(x)$ with finite m.g.f. in $(-\infty, 1)$, expectation $\mu = \tau$ and variance $\sigma^2 = \tau(1 - \tau)$. Finally set for all nonnegative x

$$q(x) := \sqrt{\tau(1-\tau)}p_1\left(x\sqrt{\tau(1-\tau)} + \tau - 1\right)$$

which for all $0 < \tau < 1$ is the density of some r.v. Y with expectation 1 and variance 1. The m.g.f. of Y is

$$E \exp \lambda Y = e \exp\left[1 - \frac{\tau}{\sqrt{\tau(1-\tau)}}\right] \exp - \left[1 - \frac{\lambda}{\sqrt{\tau(1-\tau)}}\right]^{\tau}.$$

For $\tau = 1/2$, Y has the Inverse Gaussian distribution with parameters (1, 1) and m.g.f

$$E \exp \lambda Y = e\left(\exp - [1 - 2\lambda]^{1/2}\right).$$

The variance function of the NEF generated by a stable distribution with index τ in (0, 1) writes

$$V(x) = C_\tau x^{\frac{2-\tau}{1-\tau}}$$

with

$$C_\tau := \left(\frac{1-\tau}{\tau}\right)^{\frac{2-\tau}{2(1-\tau)}}.$$

4.4.2 Compound Gamma–Poisson Distributions

We briefly characterize this compound distribution and the resulting weight Y. Let μ denote the distribution of $S_N := \sum_{i=0}^{N} \Gamma_i$ where $S_0 := 0$, N is a Poisson (p) r.v. independent of the independent family $(\Gamma_i)_{i\geq 1}$ where the Γ_i's are distributed with Gamma distribution with scale parameter $1/\lambda$ and shape parameter $-\rho$. Here

$$\rho := \frac{\gamma - 1}{\gamma}$$

$$\lambda := \rho$$

$$p := (\gamma - 1)^{-1/\gamma}$$

where we used the results in [28] p. 1516. Consider the family of distributions NEF(μ) generated by μ, which has power variance function $V(x) = x^{\gamma+1}$ defined on $\mathbb{R}^+$. The r.v. Y has distribution in NEF(μ) with expectation and variance 1. Its density is of the form

$$f_W(x) := \exp(ax + b) f(x)$$

where $f(x) := (d\mu(x)/dx)$ is the density of S_N. The values of the parameters a and b are

$$a := -1$$

$$b := -(\gamma - 1)^{-1/\gamma} \left[\left(1 - \frac{\gamma}{\gamma - 1}\right)^{\rho} - 1\right].$$

4.5 Cressie–Read Divergences, Weights and Variance Functions

For

$$\varphi_\gamma(x) := C \frac{x^\gamma - \gamma x + \gamma - 1}{\gamma(\gamma - 1)} \tag{20}$$

with $\gamma \neq 0, 1$, the convex function φ_γ satisfies $\varphi_\gamma(1) = \varphi'_\gamma(1) = 0$ and $\varphi''_\gamma(1) = C$, being therefore a divergence function; it is customary to assume that the positive constant C satisfies $C = 1$, a condition which we will not consider, still denoting this class of functions the Cressie–Read family of divergence functions. Set $\varphi_0(x) = -\log x + x - 1$ and $\varphi_1(x) = x \log x - x + 1$, the likelihood divergence function and the Kullback–Leibler one, noting that $\lim_{\gamma \to 0} \varphi_\alpha(x) = \varphi_0(x)$ and $\lim_{\gamma \to 1} \varphi_\gamma(x) = \varphi_1(x)$. The Cressie–Read family defined through (20) is the simplest system of nonnegative convex functions satisfying the requirements for a divergence function.

We prove that any Cressie–Read divergence function is the Fenchel Legendre transform of a moment-generating function of a random variable with expectation 1 and variance $1/C$ in a specific NEF, depending upon the divergence. Indeed we identify such a r.v. Y as follows: let Y be a r.v. with a cumulant-generating function. $\Lambda(t) := \log E \exp tY$ such that

$$\varphi_\gamma(x) = \Lambda^*(x) = \sup_t tx - \psi(t); \tag{21}$$

then

$$\frac{1}{\frac{d^2}{dx^2}\varphi_\gamma(x)} = \frac{1}{C} x^\alpha = V(x) \tag{22}$$

with $\alpha = 2 - \gamma$ for $x \to V(x)$ the variance function of the NEF generated by the distribution of Y. Since the differential equation $\frac{d^2}{dx^2}\varphi_\gamma(x) = Cx^{-\alpha}$ defines $\varphi_\gamma(x)$ through (20) in a unique way we have proved the one-to-one correspondence between Cressie–Read divergences and NEF's with power Variance functions.

Remark 5 Reproductible NEF's with power variance functions and power normalizing factors are infinitely divisible (see [28]); reciprocally all reproductible NEF's with power normalizing factors are infinitely divisible. The Cressie–Read family of divergences possesses, therefore, a quite peculiar property: they are the only ones which are the Legendre transform of cumulant- generating functions of reproductible infinitely divisible distributions with power normalizing constants. Reciprocally any wild empirical measure with reproductible infinitely divisible weights with power normalizing factors and with expectation 1 has LDP rate in the Cressie–Read family.

4.6 Examples

For example, the Tweedie scale of distributions defines random variables Y with expectation 1 and variance C_τ corresponding to Cressie–Read divergences with negative index $\gamma = -\tau/(1-\tau)$.

For $\gamma = -1$, the resulting divergence is

$$\varphi_{-1}(x) = \frac{1}{2}\frac{(x-1)^2}{x}$$

which is the modified χ^2 divergence (or Neyman χ^2). The associated r.v. Y has an Inverse Gaussian distribution with expectation 1 and variance 1.

For $\gamma = 2$ it holds

$$\varphi_2(x) = \frac{1}{2}(x-1)^2$$

which is the Spearman χ^2 divergence. The resulting r.v. Y has a Gaussian distribution with expectation 1 and variance 1. Note that in this case, Y is not a positive random variable.

For $\gamma = 1/2$ we get

$$\varphi_{1/2}(x) = 2\left(\sqrt{x}-1\right)^2$$

which is the Hellinger divergence. The associated random variable Y has a Compound Gamma–Poisson distribution with $\rho = -1, \lambda = -1, p = 4, a = -1$ and $b = 4$.

When $\gamma = 3/2$ the distribution of Y belongs to the NEF generated by the stable law μ on $\mathbb{R}^+$ with characteristic exponent $1/3$, hence with density the Modified Bessel-type distribution

$$f(x) = (d\mu(x)/dx) = (2\pi)^{-1}\lambda K_{1/2}\left(\lambda x^{1/2}\right)\exp\left(-px + 3\left(\lambda^2 p/4\right)^{1/3}\right)$$

where λ and p are positive and $K_{1/2}(z)$ is the modified Bessel function of order $1/2$ with argument z.

When $\gamma = 1$ then

$$\varphi_0(x) = x\log x - x + 1,$$

the Kullback–Leibler divergence function, and Y has a Poisson distribution with parameter 1. Since the rate of the corresponding LDP coincides with the rate of the LDP for the empirical distribution of the data (unconditionally), and since the variance function characterizes the distribution of the weights, this is the only wild bootstrap which is LDP efficient.

When $\gamma = 0$ then

$$\varphi_0(x) = -\log x + x - 1,$$

the Likelihood divergence and Y has an exponential with parameter 1.

The L_1 divergence function $\varphi(x) = |x-1|$ does not yield to any weighted sampling; indeed $\varphi^*(t) = t1_{(-1,1)}(t) + \infty 1_{(-1,1)^c}(t)$ which is not a cumulan- generating function.

5 Monte Carlo Minimization of a Cressie–Read Divergence Through Wild Bootstrap

Due to the preceding correspondence between the minimization problem (P) and Large Deviation rates, we propose the following procedures for the estimation of $\phi(\Omega, P)$.

Simulate nK i.i.d. random variables $Y, Y_{1,i}, Y_{2,i}, \ldots, Y_{K,i}$, $1 \leq i \leq n$ with common distribution in correspondence with the divergence function φ, namely such that

$$\varphi(x) = \Lambda^*(x)$$

for $x \in Dom\varphi$ where $\Lambda^*(x) := \sup_t tx - \Lambda(t)$ and $\Lambda(t) = \log E \exp tY$. Define

$$P_{n,K}(\Omega) := \frac{1}{K} card\left(j \in \{1, \ldots, K\} : P_{n,j}^{Wild} \in \Omega\right)$$

where

$$P_{n,j}^{Wild} := \frac{1}{n}\sum_{i=1}^{n} Y_{j,i}\delta_{x_i}$$

$1 \leq j \leq K$.

Define

$$\phi_{n,K}^{Wild}(\Omega, P) := -\frac{1}{n}\log P_{n,K}(\Omega) .$$

6 Sets of Measures for Which the Monte Carlo Minimization Technique Applies

We explore cases when

$$\phi(int(\Omega), P) = \phi(cl(\Omega), P) \tag{23}$$

in the weak topology on $\mathscr{M}$. Two conditions are derived; in the first case we make use of convexity arguments; we make use of a similar argument as used in [30], Corollary 3.1. For Ω a subset of $\mathscr{M}$ denote $cl_w((\Omega))$, resp. $int_w(\Omega)$, the weak closure (resp.) the weak interior of Ω in $\mathscr{M}$.

A convex set Ω in $\mathscr{M}$ is strongly $w-$convex if for all Q in $cl_w((\Omega))$ and each R in $int_w(\Omega)$ it holds that

$$\{\alpha Q + (1-\alpha)R; 0 < \alpha < 1\} \subset int_w(\Omega) .$$

It holds

Proposition 3 *Let $P \in \mathscr{M}_1$ and let $\Omega_1, \ldots, \Omega_J$ be subsets of $\mathscr{M}$. Set $\Omega := \Omega_1 \cup \ldots \cup \Omega_J$. Then when all Ω_j s are strongly $w-$convex and $\phi(int_w\left(\Omega_j\right), P) < \infty$ for all j, (23) holds.*

Proof For any $j = 1, \ldots, J$, fix $\varepsilon > 0$. Let $Q \in cl_w\left(\left(\Omega_j\right)\right)$ be such that

$$\phi\left(Q, P\right) < \phi\left(cl_w\left(\Omega_j\right), P\right) + \varepsilon$$

and $R \in int_w\left(\Omega_j\right)$ be such that $\phi\left(R, P\right) < \infty$. Define $Q_\alpha := \alpha Q + (1-\alpha)R$, $0 < \alpha < 1$. Then $Q_\alpha \in int_w\left(\Omega_j\right)$ and the convexity of $Q' \to \phi\left(Q', P\right)$ implies

$$\begin{aligned}\phi(int_w\left(\Omega j\right), P) &\leq \lim_{\alpha\uparrow 1}\left\{\alpha\phi\left(Q, P\right) + (1-\alpha)\phi\left(R, P\right)\right\}\\ &= \phi\left(Q, P\right) < \phi\left(cl_w\left(\left(\Omega_j\right)\right), P\right) + \varepsilon.\end{aligned}$$

Hence $\phi(int_w\left(\Omega_j\right), P) = \phi\left(cl_w\left(\left(\Omega_j\right)\right), P\right)$. Therefore, (23) holds for the finite union of the Ω_j's, as sought.

Some other class of sets $\Omega \subset \mathscr{M}$ for which (23) holds are defined as pre-images of continuous linear functions defined from $\mathscr{X}$ onto some Hausdorff topological space E. Adapting Theorem 4.1 in [30] we may state

Proposition 4 *Let $P \in \mathscr{M}_1$ and E be a real Hausdorff topological space; let $B_1 \subset B_2 \subset \ldots$ be an increasing sequence of Borel sets in supp(P) such that*

$$\lim_{m\to\infty} P(B_m) = 1.$$

Let $\Psi_m := \{Q \in \mathscr{M} : |Q|\,(B_m) = 1\}$ for all $m \in \mathbb{N}$ and $\mathscr{M}^* := \cup_m \Psi_m$. Let $T : \mathscr{M}^* \to E$ a function such that its restriction $T_{|\Psi_m}$ is linear and weakly continuous at each Q in $\mathscr{M}^*$ such that $\phi\left(Q, P\right) < \infty$ for each m. Let A be a convex set in E with $\phi\left(T^{-1}\left(intA\right), P\right) < \infty$. Then

$$\phi\left(T^{-1}\left(intA\right), P\right) = \phi\left(T^{-1}\left(clA\right), P\right). \tag{24}$$

Proof It proceeds following nearly verbatim the Proof of Theorem 4.1 in [30]. Convexity arguments similar to the one in the Proof of Proposition 3 provide a version of (24) for sets $T^{-1}_{|\Psi_m}(A)$. Making use of Theorem 2, which substitutes Theorem 3.1 in [30] concludes the proof.

7 A Simple Convergence Result and Some Perspectives

All estimators of $\phi\left(\Omega, P\right)$ considered in this paper converge strongly to $\phi\left(\Omega, P\right)$ as n tends to infinity, as does K. Indeed going back to the general setting presented in Sect. 1, for fixed n it clearly holds that

$$\lim_{K\to\infty} \frac{1}{K} card\left(i : X_{n,i} \in \Omega\right) = \Pr\left(X_n \in \Omega\right)$$

a.s.

When

$$\lim_{n\to\infty} \frac{1}{n} \log \Pr\left(X_n \in \Omega\right) = -\Phi\left(\Omega\right)$$

it follows that

$$\lim_{n\to\infty} \lim_{K\to\infty} \frac{1}{n} \log f_{n,K} = -\Phi\left(\Omega\right) \quad \text{a.s.}$$

as sought. Since the estimators of $KL\left(\Omega, P\right)$ and $KL_m\left(\Omega, P\right)$ considered in Sect. 3, as well as the estimators of $\phi_\gamma\left(\Omega, P\right)$ considered in Sect. 5 are obtained through continuous transformations of the former estimates, all estimators considered in the present article converge strongly to their respective limits as K tends to infinity and n tends to infinity. This leaves a large field of investigations wide open, such as the choice of some sequence $K = K_n$ which would lead to a single limit procedure. Also the resulting rate of convergence of these estimators as well as their distributional limit would be of interest.

Also Importance Sampling (IS) techniques should be investigated in order to reduce the calculation burden caused by the fact that any of the weighted empirical measures considered in this article would visit the set Ω quite rarely, if P does not belong to Ω. The hit rate can be increased substantially using some ad hoc modification of the weights, resulting from an IS strategy.

Once estimated the minimum value of the divergence, one may be interested in the identification of the measures Q which achieve this minimum in Ω. Dichotomous methods can be used, iterating the evaluation of the minimum divergence between P and subsets of Ω where the global infimum on Ω coincides with the local ones, leaving apart the subsets where they do not coincide, and iterating this routine.

References

1. Broniatowski, M., Keziou, A.: Divergences and duality for estimation and test under moment condition models. J. Stat. Plann. Inference **142**(9), 2554–2573 (2012)
2. Broniatowski, M., Decurninge, A.: Estimation for models defined by conditions on their L-moments. IEEE Trans. Inf. Theory. doi:10.1109/TIT.2016.2586085
3. Ruschendorf, L.: Projections of probability measures. Statistics **18**(1), 123–129 (1987)
4. Csiszár, I.: Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. (German) Magyar Tud. Akad. Mat. Kutató Int. Közl. 8, 85–108 (1963)
5. Csiszár, I., Gamboa, F., Gassiat, E.: MEM pixel correlated solutions for generalized moment and interpolation problems. IEEE Trans. Inf. Theory **45**(7), 2253–2270 (1999)
6. Csiszár, I.: Information-type measures of difference of probability distributions and indirect observations. Studia Sci. Math. Hungar. **2**, 299–318 (1967)
7. Csiszár, I.: On topology properties of f-divergences. Studia Sci. Math. Hungar. **2**, 329–339 (1967)

8. Liese, F., Vajda, I.: Convex statistical distances. With German, French and Russian summaries. Teubner-Texte zur Mathematik [Teubner Texts in Mathematics], 95. BSB B. G. Teubner Verlagsgesellschaft, pp. 224. Leipzig (1987). ISBN: 3-322-00428-7
9. Broniatowski, M., Keziou, A.: Minimization of φ-divergences on sets of signed measures. Studia Sci. Math. Hungar. **43**(4), 403–442 (2006)
10. Cressie, N., Read, T.R.C.: Multinomial goodness-of-fit tests. J. Roy. Stat. Soc. Ser. B **46**(3), 440–464 (1984)
11. Broniatowski, M., Keziou, A.: Parametric estimation and tests through divergences and the duality technique. J. Multivar. Anal. **100**(1), 16–36 (2009)
12. Keziou, A.: Dual representation of φ-divergences and applications. C. R. Math. Acad. Sci. Paris **336**(10), 857–862 (2003)
13. Owen, A.: Empirical likelihood ratio confidence regions. Ann. Stat. **18**(1), 90–120 (1990)
14. Bertail, P.: Empirical likelihood in some semiparametric models. Bernoulli **12**(2), 299–331 (2006)
15. Jiménez, R., Shao, Y.: On robustness and efficiency of minimum divergence estimators. Test **10**(2), 241–248 (2001)
16. Beran, R.: Minimum Hellinger distance estimates for parametric models. Ann. Stat. **5**(3), 445–463 (1977)
17. Lindsay, B.G.: Efficiency versus robustness: the case for minimum Hellinger distance and related methods. Ann. Stat. **22**(2), 1081–1114 (1994)
18. Rockafellar, R.: Tyrrell Convex Analysis, Princeton Mathematical Series, vol. 28, pp. xviii+451. Princeton University Press, Princeton, NJ (1970)
19. Barndorff-Nielsen, O.: Information and Exponential Families in Statistical Theory. Wiley Series in Probability and Mathematical Statistics, pp. ix+238. Wiley, Ltd., Chichester (1978)
20. Trashorras, J., Wintenberger, O.: Large deviations for bootstrapped empirical measures. Bernoulli **20**(4), 1845–1878 (2014)
21. Najim, J.: A Cramer type theorem for weighted random variables. Electron. J. Probab. **7** (2002)
22. Broniatowski, M. Weighted sampling, maximum likelihood and minimum divergence estimators. In: Geometric science of information, Lecture Notes in Computer Science, vol. 8085, pp. 467–478. Springer, Heidelberg (2013)
23. Barbe, P., Bertail, P.: The Weighted Bootstrap, Lecture Notes in Statistics. Springer, New York (1995)
24. Letac, G., Mora, M.: Natural real exponential families with cubic variance functions. Ann. Stat. **18**(1), 1–37 (1990)
25. Tweedie, M.C.K.: Functions of a statistical variate with given means, with special reference to Laplacian distributions. Proc. Camb. Philos. Soc. **43**, 41–49 (1947)
26. Jørgensen, B.: Exponential dispersion models. With discussion and a reply by the author. J. Roy. Stat. Soc. Ser. B **49**(2), 127–162 (1987)
27. Morris, C.N.: Natural exponential families with quadratic variance functions. Ann. Stat. **10**(1), 65–80 (1982)
28. Bar-Lev, S.K., Enis, P.: Reproducibility and natural exponential families with power variance functions. Ann. Stat. **14**(4), 1507–1522 (1986)
29. Feller, W.: An Introduction to Probability Theory and its Applications, Vol. II, 2nd edn., pp. xxiv+669. Wiley, Inc., New York (1971)
30. Groeneboom, P., Oosterhoff, J., Ruymgaart, F.H.: Large deviation theorems for empirical probability measures. Ann. Probab. **7**(4), 553–586 (1979)
31. Read, T.R.C., Cressie, N.A.C.: Goodness-of-fit Statistics for Discrete Multivariate Data. Springer Series in Statistics, pp. xii+211. Springer, New York (1988). ISBN: 0-387-96682-X
32. Mason, D.M., Newton, M.A.: A rank statistic approach to the consistency of a general bootstrap. Ann. Stat. **20**, 1611–1624 (1992)

33. Hoadley, A.B.: On the probability of large deviations of functions of several empirical cdf's. Ann. Math. Stat. **38**, 360–381 (1967)
34. Withers, C.S., Nadarajah, S.: On the compound Poisson-gamma distribution. Kybernetika (Prague) **47**(1), 15–37 (2011)
35. Several Applications of Divergence Criteria in Continuous Families, to appear Kybernetika (2012)

Asymptotic Analysis of Iterated 1-Step Huber-Skip M-Estimators with Varying Cut-Offs

Xiyu Jiao and Bent Nielsen

Abstract We consider outlier detection algorithms for time series regression based on iterated 1-step Huber-skip M-estimators. This paper analyses the role of varying cut-offs in such algorithms. The argument involves an asymptotic theory for a new class of weighted and marked empirical processes allowing for estimation errors of the scale and the regression coefficient.

Keywords The iterated 1-step Huber-skip M-estimator · Tightness · A fixed point · Poisson approximation to gauge · Weighted and marked empirical processes

1 Introduction

We consider outlier detection methods that are based on iterated 1-step Huber-skip M-estimators for linear regression models with regressors that are stationary or deterministically or stochastically trending. Each 1-step estimator relies on a cut-off value when classifying observations as outliers or not. In this paper, we allow the cut-off value to vary with sample size and iteration step. To analyze this asymptotically, we generalize some recent results for residual empirical processes, which allow for variation in location, scale and quantile. The model is a linear regression

$$y_i = x_i'\beta + \varepsilon_i, \quad i = 1, 2, \ldots, n, \tag{1}$$

where ε_i/σ are independent of $\mathscr{F}_{i-1} = \sigma(x_1, \ldots, x_i, \varepsilon_1, \ldots, \varepsilon_{i-1})$ with the common density f. Outliers are pairs of observations that do not conform with the model.

X. Jiao (✉)
Department of Economics, University of Oxford and Mansfield College,
Oxford OX1 3TF, UK
e-mail: xiyu.jiao@mansfield.ox.ac.uk

B. Nielsen
Department of Economics, University of Oxford and Nuffield College,
Oxford OX1 1NF, UK
e-mail: bent.nielsen@nuffield.ox.ac.uk

J. Antoch et al. (eds.), *Analytical Methods in Statistics*, Springer Proceedings
in Mathematics & Statistics 193, DOI 10.1007/978-3-319-51313-3_2

Iterated 1-step Huber-skip M-estimators mimic the Huber [14] skip estimator, which has criterion function $\rho(t) = \min(t^2, c^2)/2$ as opposed to the Huber estimator with criterion function $\rho(t) = t^2/2$ for $|t| \leq c$ and $\rho(t) = c|t| - c^2/2$ otherwise, see also [8, p. 104], [19, p. 175]. The 1-step Huber-skip M-estimator starts from an initial estimator $(\widetilde{\beta}, \widetilde{\sigma}^2)$. This is used to decide which observations are outlying through

$$v_i = 1_{(|y_i - x_i'\widetilde{\beta}| \leq \widetilde{\sigma} c)}, \tag{2}$$

where the choice of the cut-off c is related to the known reference density f. For those observations that are not outlying, we run a least squares regression and get the 1-step Huber-skip estimator

$$\widehat{\beta} = \left(\sum_{i=1}^n x_i x_i' v_i\right)^{-1} \left(\sum_{i=1}^n x_i y_i v_i\right), \tag{3}$$

$$\widehat{\sigma}^2 = \varsigma^{-2} \left(\sum_{i=1}^n v_i\right)^{-1} \left\{\sum_{i=1}^n (y_i - x_i'\widehat{\beta})^2 v_i\right\}, \tag{4}$$

where ς^2 is the consistency factor as in (8). This step can be iterated. The iteration may be initiated by a robust estimator. More simply we get the Robustified Least Squares and the Impulse Indicator Saturation starting with the full or split sample least squares. The latter algorithm was introduced in the empirical work of US food expenditure by Hendry, see [9, 10].

Outlier detection algorithms have a positive probability to find outliers even when, in fact, the data generation process has no outliers. We evaluate the performance of such algorithms by the concept of a gauge, which is the expected retention rate of falsely discovered outliers. This is a measure of type I error and it gives us an indirect way of choosing the cut-off c. It is defined as follows. The algorithms assign stochastic indicators v_i to all observations such as in (2) so that $v_i = 0$ when observation i is declared as an outlier, otherwise $v_i = 1$. When the model has no contamination, the sample and population gauge are

$$\widehat{\gamma} = \frac{1}{n} \sum_{i=1}^n (1 - v_i), \qquad \mathsf{E}\widehat{\gamma} = \frac{1}{n} \sum_{i=1}^n \mathsf{E}(1 - v_i). \tag{5}$$

Hoover and Perez [13] originally introduced the idea of a gauge in a simulation study of general-to-specific variable selection algorithms. The concept of a gauge was formally proposed by Hendry and Santos [12] as the expected retention rate of irrelevant regressors in the context of model selection algorithms. Comprehensive simulation studies on the gauge for the model selection algorithm Autometrics are presented in [6, 10]. An asymptotic analysis for the gauge of some outlier detection algorithms is presented in [18].

One-step estimators have been considered before in [2, 23]. The 1-step Huber-skip estimator was studied in [25]. Asymptotic distribution theory has been derived for the location model in [11] and for the time series regression [15]. Iteration was investigated in [16]. An asymptotic expansion for the sample gauge was established in [18]. All these asymptotic analyses are restricted to the situation where the cut-off and the number of iterations are not both increasing.

The purpose of this paper is to build an asymptotic theory which can explore how variation in the cut-off affects the iterated 1-step Huber-skip M-estimator. In particular, we prove the tightness and fixed point theorems for the iterated 1-step M-estimator with the varying cut-off. Moreover, this paper demonstrates an asymptotic Poisson distribution to the gauge in a situation where the cut-off increases with the sample size while the number of iterations also increases.

The argument involves a theory for a new class of weighted and marked empirical processes. This is defined from the generalized empirical distribution function

$$\widehat{\mathsf{F}}_n^{g,p}(a, b, c) = \frac{1}{n}\sum_{i=1}^{n} g_{in}\varepsilon_i^p 1_{(\varepsilon_i \le \sigma c + n^{-1/2}ac + x_{in}'b)}, \tag{6}$$

where the weights g_{in} are combinations of the normalized $\mathscr{F}_{i-1}$ measurable regressors x_{in} and ε_i^p are the $\mathscr{F}_i$ adapted marks, while a, b represent the normalized estimation errors for σ, β. When $p = 0$ the mark is unity and we get the weighted empirical distribution function considered by for instance [20]. Processes of the type $n^{-1/2}\sum_{i=1}^{n} \varepsilon_i 1_{(x_i \le c)}$ are called marked processes, see [20, p. 43], but are not special cases of the weighted and marked empirical distribution functions.

We derive asymptotic expansions that are uniform in a, b, c and allow for a near $n^{1/4}$ inefficiency in the estimation uncertainties a, b. This generalizes results by Koul and Ossiander, see [20–22], who allowed unbounded weights g_{in} but no marks ε_i^p. They used a truncation argument for $\mathscr{F}_{i-1}$ measurable weights g_{in}. This together with the boundedness of the $\mathscr{F}_i$ measurable indicator function meant that they could apply the Freedman [7] exponential inequality for bounded martingales. Here, we use the iterated martingale inequality of [18] reported as Lemma 3 in the appendix. This is based on the Bercu and Touati [1] exponential inequality for unbounded martingales, so that we can avoid the truncation argument and more easily allow the $\mathscr{F}_i$ measurable product of the mark and indicator to be unbounded. The result also generalizes [15, 18] who did not allow joint variation of all of a, b, c.

The outline of this paper is the following. We first review the model and iterated 1-step Huber-skip M-estimator algorithm in Sect. 2. Then, the main results follow in Sect. 3. Section 4 provides theory for the weighted and marked empirical process with proofs in Appendix 1, 2, and 3. Proofs of the main theorems in Sect. 3 follow in Appendix 4.

2 Model and Outlier Detection Algorithms

The regression model with some notations is described first. We review the iterated 1-step Huber-skip M-estimators including the Robustified Least Squares and the Impulse Indicator Saturation.

2.1 Model

Suppose we have data (y_i, x_i), $i = 1, 2, \ldots, n$, where y_i is univariate and x_i is multivariate with dimension $\dim x$. Assume the data satisfies the regression equation

$$y_i = x_i'\beta + \varepsilon_i, \quad i = 1, 2, \ldots, n.$$

This setting can represent both classical regression and time series models. Moreover, regressors x_i can be a deterministic or stochastic trend. Innovations ε_i are independent of the filtration $\mathscr{F}_{i-1}$ generated by $(x_1, \ldots, x_i, \varepsilon_1, \ldots, \varepsilon_{i-1})$, and are identically distributed with scale σ so that ε_i/σ has the known density f and distribution function $\mathsf{F}(c) = \mathsf{P}(\varepsilon_i/\sigma \leq c)$. In practice, the innovation distribution, characterized by f, F, will often be assumed to be standard normal or at least symmetric. Outlier detection algorithms use absolute residuals and then calculate robust least squares estimators from the non-outlying sample. This implicitly assumes symmetry, while non-symmetry leads to bias forms. We assume symmetry when analyzing the iterated 1-step Huber-skip M-estimator algorithm in Sect. 3, but not for the general empirical process results in Sect. 4.

For the absolute error $|\varepsilon_i|/\sigma$ we denote the density by g and the distribution function by $\mathsf{G}(c) = \mathsf{P}(|\varepsilon_i|/\sigma \leq c)$ for $c > 0$. Here we use c as notation for the quantile of the distribution $\mathsf{G}(c)$. In the course of the analysis this will be linked to the cut-off of the 1-step estimator in (3) and the argument of the weighted and marked empirical distribution function in (6). Now, with a symmetry assumption, $\mathsf{G}(c) = 2\mathsf{F}(c) - 1$ and $\mathsf{g}(c) = 2\mathsf{f}(c)$. Define $\psi = \mathsf{G}(c)$ so the probability of exceeding the cut-off c is $\gamma = 1 - \psi$. Suppose the k-th moment of the density f exists, then introduce

$$\tau_k = \int_{-\infty}^{\infty} u^k \mathsf{f}(u) du, \qquad \tau_k^c = \int_{-c}^{c} u^k \mathsf{f}(u) du. \tag{7}$$

Thus $\tau_0^c = \psi$, $\tau_2 = 1$ while $\tau_k = \tau_k^c = 0$ for odd k when assuming symmetry. Define the conditional variance of ε_i/σ given $(|\varepsilon_i|/\sigma \leq c)$ as

$$\varsigma_c^2 = \frac{\tau_2^c}{\psi} = \frac{\int_{-c}^{c} u^2 \mathsf{f}(u) du}{\mathsf{P}(|\varepsilon_i| \leq \sigma c)}. \tag{8}$$

This will be used as a bias correction factor for the variance estimate computed from the selected non-outlying sample. For a standard normal reference distribution, we have $\tau_2^c = \psi - 2c\mathsf{f}(c)$, $\tau_4^c = 3\psi - 2c(c^2+3)\mathsf{f}(c)$ and $\tau_4 = 3$.

2.2 The Iterated 1-Step Huber-Skip M-Estimator Algorithm

We first define the iterated 1-step Huber-skip M-estimator algorithm. Specific examples include the Robustified Least Squares and the Impulse Indicator Saturation.

Algorithm 1 ***Iterated 1-step Huber-skip M-estimator***. *Choose a cut-off $c > 0$.*

1. *Choose initial estimators $\widehat{\beta}_c^{(0)}$, $(\widehat{\sigma}_c^{(0)})^2$ and let $m = 0$.*
2. *Define indicator variables for selecting non-outlying observations*

$$v_{i,c}^{(m)} = 1_{(|y_i - x_i'\widehat{\beta}_c^{(m)}| \le \widehat{\sigma}_c^{(m)} c)}. \tag{9}$$

3. *Compute least squares estimators*

$$\widehat{\beta}_c^{(m+1)} = \left(\sum_{i=1}^n x_i x_i' v_{i,c}^{(m)}\right)^{-1} \left(\sum_{i=1}^n x_i y_i v_{i,c}^{(m)}\right), \tag{10}$$

$$(\widehat{\sigma}_c^{(m+1)})^2 = \varsigma_c^{-2}\left(\sum_{i=1}^n v_{i,c}^{(m)}\right)^{-1} \left\{\sum_{i=1}^n (y_i - x_i'\widehat{\beta}_c^{(m+1)})^2 v_{i,c}^{(m)}\right\}. \tag{11}$$

4. *Let $m = m + 1$ and repeat 2 and 3.*

In Sect. 3 we show how to choose the cut-off c indirectly from the gauge defined in (5). The algorithm could start with a robust estimator, while the Robustified Least Squares is initiated using the full sample least squares. The latter is not robust with respect to high leverage points in cross section data. Leverage points seem to be less of a problem in time series models when lagged variables are included as regressors.

Another example is the Impulse Indicator Saturation which was initially proposed in the empirical work [9]. The algorithm was studied comprehensively in [10, 11]. The idea is to divide full sample into two sub-samples and use regression estimates calculated from each sub-sample to detect outliers in the other sub-sample.

Algorithm 2 ***Impulse Indicator Saturation***. *Choose a cut-off $c > 0$.*

1.1. Split full sample into two sets $\mathscr{I}_j, j = 1, 2$ of n_j observations where $\sum_{j=1}^2 n_j = n$.
1.2. Calculate least squares estimators based upon each sub-sample $\mathscr{I}_j$ for $j = 1, 2$

$$\widehat{\beta}_j = \left(\sum_{i\in\mathscr{I}_j} x_i x_i'\right)^{-1} \left(\sum_{i\in\mathscr{I}_j} x_i y_i\right), \quad \widehat{\sigma}_j^2 = \frac{1}{n_j}\sum_{i\in\mathscr{I}_j}(y_i - x_i'\widehat{\beta}_j)^2. \tag{12}$$

1.3. Define the initial indicator variables for selecting non-outlying observations

$$v_{i,c}^{(-1)} = 1_{(i\in\mathscr{I}_1)}1_{(|y_i - x_i'\widehat{\beta}_2|\leq\widehat{\sigma}_2 c)} + 1_{(i\in\mathscr{I}_2)}1_{(|y_i - x_i'\widehat{\beta}_1|\leq\widehat{\sigma}_1 c)}. \tag{13}$$

1.4. Compute $\widehat{\beta}_c^{(0)}$, $(\widehat{\sigma}_c^{(0)})^2$ *using (10) and (11) with* $m = -1$, *and then let* $m = 0$.

2. Follow the step 2,3,4 in Algorithm 1.

The Impulse Indicator Saturation is possibly more robust than the Robustified Least Squares when we have prior knowledge that outliers are located in a particular subset of the whole sample. The choice of the initial sets $\mathscr{I}_1$ and $\mathscr{I}_2$ should be iterated since the location of contaminated observations is unknown in most practical situations, see [6].

3 The Main Results

We start by listing the assumptions. Then follows the new tightness and fixed point result for the iterated estimator defined in Algorithm 1. Finally the gauge of the iterated estimator is analyzed. The result is uniform in the cut-off value, which generalizes [15, 16] which set the threshold fixed. This allows us to analyze the gauge of the iterated estimator when the cut-off value is drifting.

3.1 Assumptions

We list the sufficient assumptions for asymptotic theory of iterated 1-step Huber-skip M-estimators. These assumptions are somewhat stronger than they need to be. In Sect. 4 on the one-sided empirical process, we will introduce some weaker assumptions. For instance, we will then abandon the symmetry assumption of f.

Innovations ε_i and regressors x_i must satisfy some moment conditions so as to carry out asymptotic analysis. Regressors x_i can be temporally dependent and trending deterministically or stochastically. We therefore need a normalisation matrix N that allows for different behaviour of the components of the regressor vector x_i. In the case of a stationary regressor we need a standard $n^{-1/2}$ normalisation so that N must be proportional to the identity matrix of the same dimension as x_i, that is $N = n^{-1/2}I_{\dim x}$. Likewise, if x_i is a random walk we have $N = n^{-1}I_{\dim x}$. If the regressors are unbalanced as in $x_i = (1, i)'$ we can choose $N = \text{diag}(n^{-1/2}, n^{-3/2})$.

Assumption 1 Let $\mathscr{F}_i$ be an increasing sequence of σ-fields so ε_{i-1} and x_i are $\mathscr{F}_{i-1}$ measurable and ε_i is independent of $\mathscr{F}_{i-1}$. Let ε_i/σ have a symmetric, continuously differentiable density f which is positive on $\mathbf{R}$. For some values of κ, η such that $0 \leq \kappa < \eta \leq 1/4$, choose an integer $r \geq 2$ so

$$2^{r-1} \geq 1 + (1/4 + \kappa - \eta)(1 + \dim x). \tag{14}$$

Let $q = 1 + 2^{r+1}$. Denote $c_0 > 0$ as a finite number. Suppose

(i) the density f satisfies

  (a) $u^q \mathsf{f}(u)$, $|u^{q+1}\dot{\mathsf{f}}(u)|$ are decreasing for large u;
  (b) $\mathsf{f}(u_n - n^{-1/4}A)/\mathsf{f}(u_n) = \mathrm{O}(1)$ as $n \to \infty$ for some $A > 0$ and all sequences $u_n \to \infty$ so $u_n = \mathrm{o}(n^{1/4})$;
  (c) $\mathsf{f}(u)/[u\{1 - \mathsf{F}(u)\}] = \mathrm{O}(1)$ for $u \to \infty$;

(ii) the regressors x_i satisfy

  (a) $\Sigma_n = \sum_{i=1}^n N' x_i x_i' N \overset{\mathsf{P}}{\to} \Sigma \overset{a.s.}{>} 0$;
  (b) $\max_{1 \le i \le n} |n^{1/2-\kappa} N' x_i| = \mathrm{O}_{\mathsf{P}}(1)$;
  (c) $n^{-1} \mathsf{E} \sum_{i=1}^n |n^{1/2} N' x_i|^q = \mathrm{O}(1)$;

(iii) the initial estimator $(\widetilde{\beta}, \widetilde{\sigma}^2)$ satisfies

  (a) $N^{-1}(\widetilde{\beta} - \beta) = \mathrm{O}_{\mathsf{P}}(n^{1/4-\eta})$;
  (b) $n^{1/2}(\widetilde{\sigma}^2 - \sigma^2) = \mathrm{O}_{\mathsf{P}}(n^{1/4-\eta})$.

There is a trade-off between κ, η, the dimension dim x and the required number of moments r, see [17, Remark 3.1]. The conditions (i), (ii) are satisfied in a range of situations. In particular, condition (ia) is satisfied by the normal and t distribution, see [17, Example 3.1]; condition (ib, ic) is satisfied by the normal, see [18, Remark 2]; condition (ii) is satisfied by stationary, random walk and deterministically trending regressors, see [17, Example 3.2]. Condition (iii) allows the standardized estimation errors to diverge at a rate of $n^{1/4-\eta}$ rather than being bounded in probability. In particular, $\eta = 1/4$ can be chosen for estimators with standard convergence rates.

3.2 Properties of the Iterated Estimators

The first result is a stochastic expansion of the 1-step Huber-skip M-estimator in terms of the original estimator, a kernel, and a small remainder term.

Theorem 1 *Consider the iterated 1-step Huber-skip M-estimator in Algorithm 1. Suppose Assumption 1(ia, ii) holds, and that* $N^{-1}(\widehat{\beta}_c^{(m)} - \beta)$, $n^{1/2}(\widehat{\sigma}_c^{(m)} - \sigma)$ *are* $\mathrm{O}_{\mathsf{P}}(1)$. *Then uniformly in* $c \in [c_0, \infty)$ *and as* $n \to \infty$

$$N^{-1}(\widehat{\beta}_c^{(m+1)} - \beta) = \frac{2c\mathsf{f}(c)}{\psi} N^{-1}(\widehat{\beta}_c^{(m)} - \beta) + (\psi \Sigma)^{-1} \sum_{i=1}^n N' x_i \varepsilon_i 1_{(|\varepsilon_i| \le \sigma c)} + \mathrm{o}_{\mathsf{P}}(1),$$

$$n^{1/2}(\widehat{\sigma}_c^{(m+1)} - \sigma) = \frac{c(c^2 - \varsigma_c^2)\mathsf{f}(c)}{\tau_2^c} n^{1/2}(\widehat{\sigma}_c^{(m)} - \sigma) + \frac{1}{2\sigma\tau_2^c} n^{-1/2} \sum_{i=1}^n (\varepsilon_i^2 - \varsigma_c^2 \sigma^2) 1_{(|\varepsilon_i| \le \sigma c)} + \mathrm{o}_{\mathsf{P}}(1).$$

Theorem 1 shows that the updated estimation error for β depends on the previous estimation error for β, but not on the estimation uncertainty for σ. The estimation error for σ has a similar property. This is a consequence of symmetry imposed on the density f. More complex situations can also be analyzed where the reference distribution f is non-symmetric and the cut-off c is chosen in a matching way, see [15]. The proof uses the empirical process theory in Sect. 4.

The next result shows that the iterated estimator is tight in iteration $m \in [0, \infty)$ and in the cut-off value $c \in [c_0, \infty)$. This builds on [16].

Theorem 2 *Consider the iterated 1-step Huber-skip M-estimator in Algorithm 1. Suppose Assumption 1(ia, ii, iii) holds with $\eta = 1/4$. Then as $n \to \infty$*

$$\sup_{0 \leq m < \infty} \sup_{c_0 \leq c < \infty} |N^{-1}(\widehat{\beta}_c^{(m)} - \beta)| + |n^{1/2}(\widehat{\sigma}_c^{(m)} - \sigma)| = \mathsf{O}_\mathsf{P}(1).$$

Assumption 1(iii) with $\eta = 1/4$ corresponds to a standard convergence rate for the initial estimator. Theorem 1 provides the 1-step relationship between the updated estimator and the original estimator. Since $\sup_{c_0 \leq c < \infty} |2c\mathsf{f}(c)/\psi| < 1$ and $\sup_{c_0 \leq c < \infty} |c(c^2 - \varsigma_c^2)\mathsf{f}(c)/\tau_2^c| < 1$ implied by Assumption 1(ia), see [16, Theorem 3.5], a geometric argument and mathematical induction are used to show tightness.

The fixed point result can now be shown. Initially the tight estimator is assumed available. This is iterated through the 1-step equation presented in Theorem 1.

Theorem 3 *Consider the iterated 1-step Huber-skip M-estimator in Algorithm 1. Suppose Assumption 1(ia, ii, iii) holds with $\eta = 1/4$. Then for all $\varepsilon, \delta > 0$ a pair $m_0, n_0 > 0$ exists, so for all $m > m_0$ and $n > n_0$*

$$\mathsf{P}\left\{\sup_{c_0 \leq c < \infty} |N^{-1}(\widehat{\beta}_c^{(m)} - \widehat{\beta}_c^{*})| + |n^{1/2}(\widehat{\sigma}_c^{(m)} - \widehat{\sigma}_c^{*})| > \delta\right\} < \varepsilon,$$

where

$$N^{-1}(\widehat{\beta}_c^{*} - \beta) = \frac{1}{\psi - 2c\mathsf{f}(c)} \Sigma^{-1} \sum_{i=1}^{n} N' x_i \varepsilon_i 1_{(|\varepsilon_i| \leq \sigma c)},$$

$$n^{1/2}(\widehat{\sigma}_c^{*} - \sigma) = \frac{1}{2\sigma\{\tau_2^c - c(c^2 - \varsigma_c^2)\mathsf{f}(c)\}} n^{-1/2} \sum_{i=1}^{n} (\varepsilon_i^2 - \varsigma_c^2 \sigma^2) 1_{(|\varepsilon_i| \leq \sigma c)}.$$

Based on Theorem 2, if the initial estimator is bounded in a large compact set with large probability, then any iterated estimator takes values in the same compact set no matter what value of the cut-off c is chosen in the interval $[c_0, \infty)$. The proof of Theorem 3 is to further argue the deviation between the m-fold iterated estimator and the fixed point is the sum of two terms vanishing exponentially and in probability respectively when m and n are sufficiently large.

The iterated 1-step Huber-skip M-estimator can be seen as a special case of iteratively reweighted least squares with binary weights. Dollinger and Staudte [5] applied an influence function argument to demonstrate convergence of iteratively reweighted least squares with smooth weights. Even if the spirit is similar, our proof is different due to binary weights. The idea of iterating 1-step estimator can also be found in [4], which analyzed the first order autoregression with infinite variance.

3.3 Properties of the Gauge

Johansen and Nielsen [18] proved the Poisson approximation to the gauge for the finite step Huber-skip M-estimator. But the iterated result was not established, since they did not have the empirical process theory which investigates the varying quantile c and estimation errors for β and σ. This paper shows the Poisson approximation to the gauge for the iterated 1-step Huber-skip M-estimator.

A Poisson exceedence theory arises in the scenario where the cut-off value c is set to allow the fixed number λ of outliers regardless of the sample size n. For some $\lambda > 0$, the cut-off value c_n is set so as to let

$$n\mathsf{P}(|\varepsilon_i| > \sigma c_n) = \lambda. \tag{15}$$

Notice that $c_n \to \infty$ as $n \to \infty$. Define $v_{i,c_n}^{(m)}$, $\widehat{\beta}_{c_n}^{(m+1)}$, $(\widehat{\sigma}_{c_n}^{(m+1)})^2$ by replacing c by c_n in expressions (9–11). The corresponding sample gauge is

$$\widehat{\gamma}_{c_n}^{(m)} = \frac{1}{n}\sum_{i=1}^{n}(1 - v_{i,c_n}^{(m)}) = \frac{1}{n}\sum_{i=1}^{n} 1_{(|y_i - x_i'\widehat{\beta}_{c_n}^{(m)}| > \widehat{\sigma}_{c_n}^{(m)} c_n)}. \tag{16}$$

Theorems 2 and 3 shows that any iterated estimator is tight, so lower and upper bounds can be found for the indicators appearing in the gauge. By exploring these bounds, the following Poisson limit theorem arises.

Theorem 4 *Consider the iterated 1-step Huber-skip M-estimator in Algorithm 1. Let c_n be defined from (15). Suppose Assumption 1 holds with $\eta = 1/4$. Then for all $0 \le m < \infty$ and as $n \to \infty$, the sample gauge in (16) satisfies*

$$n\widehat{\gamma}_{c_n}^{(m)} \overset{\mathsf{D}}{\to} \text{Poisson}(\lambda).$$

Table 1 assumes that ε_i/σ follows a standard normal distribution. For a given λ, the cut-off in (15) satisfies $c_n = \Phi^{-1}\{1 - \lambda/(2n)\}$. Cut-off values are shown for $n = 100, 200$. The Poisson approximation gives the probability of finding at most x outliers. There is an increase from 62 to 90% for the probability of detecting at most $x = \lambda$ outliers as λ declines from 5 to 0.1. The reason is due to the left skewness of the Poisson distribution. In particular, we focus on the case where $\lambda = 1$ and $n = 100$. The cut-off is $c_n = 2.58$ and the probability to find at most 1, 2 outliers are 0.74, 0.92. This means it regularly finds 2 outliers when there are none.

Table 1 The probability of detecting at most x outliers approximated by a Poisson distribution for a given λ, and the cut-off $c_n = \Phi^{-1}\{1 - \lambda/(2n)\}$ for $n = 100, 200$

			x					
λ	c_{100}	c_{200}	0	1	2	3	4	5
5	1.960	2.241	0.01	0.04	0.12	0.27	0.44	0.62
1	2.576	2.807	0.37	0.74	0.92	0.98	1.00	
0.5	2.807	3.023	0.61	0.91	0.98	1.00		
0.25	3.023	3.227	0.78	0.97	1.00			
0.1	3.291	3.481	0.90	1.00				

The Robustified Least Squares and Impulse Indicator Saturation are special versions of iterated 1-step Huber-skip M-estimators with different starting points. Their initial points do not depend on the cut-off, and thus satisfy the tightness property. Therefore, Theorems 1–4 apply for these algorithms.

4 Weighted and Marked Empirical Process

Consider the weighted and marked empirical distribution function

$$\widehat{\mathsf{F}}_n^{g,p}(a, b, c) = \frac{1}{n}\sum_{i=1}^{n} g_{in}\varepsilon_i^p 1_{(\varepsilon_i \leq \sigma c + n^{-1/2}ac + x_{in}'b)},$$

with $\mathscr{F}_{i-1}$ adapted weights g_{in} and $\mathscr{F}_i$ measurable marks ε_i^p. Let $a \in \mathbf{R}$, $b \in \mathbf{R}^{\dim x}$ represent estimation errors $\widetilde{a} = n^{1/2}(\widetilde{\sigma} - \sigma)$, $\widetilde{b} = N^{-1}(\widetilde{\beta} - \beta)$, while $c \in \mathbf{R}$ is the quantile. Define normalized regressors $x_{in} = N'x_i$ so that $\sum_{i=1}^n x_{in}x_{in}'$ converges. For example, $N = n^{-1/2}I_{\dim x}$ if $\{x_i\}_{i=1}^n$ is stationary, while $N = n^{-1}I_{\dim x}$ for a random walk. Our interest focuses on weights g_{in} given as either of 1, $n^{1/2}N'x_i$, $nN'x_ix_i'N$ and p as either of 0, 1, 2. To form the empirical process, introduce the compensator

$$\overline{\mathsf{F}}_n^{g,p}(a, b, c) = \frac{1}{n}\sum_{i=1}^{n} g_{in}\mathsf{E}_{i-1}\varepsilon_i^p 1_{(\varepsilon_i \leq \sigma c + n^{-1/2}ac + x_{in}'b)}, \tag{17}$$

where $\mathsf{E}_{i-1}(\cdot) = \mathsf{E}(\cdot|\mathscr{F}_{i-1})$. Note that $\overline{\mathsf{F}}_n^{1,0}(0, 0, c) = \mathsf{F}(c) = \mathsf{P}(\varepsilon_i \leq \sigma c)$.

We embed these processes in the space $D[0, 1]$ of processes that are continuous from the right and with limits of left, where the space is endowed with the Skorokhod metric. We do this as follows. The indicator $1_{(\varepsilon_i \leq c)}$ and the distribution function $\mathsf{F}(c)$ can be defined as 0 or 1 when c takes the values $-\infty$ and ∞ respectively. We can then define quantiles $c_\psi = \mathsf{F}^{-1}(\psi)$ for $0 \leq \psi \leq 1$. Correspondingly we can continuously extend the definition of the weighted and marked empirical distribution function and its compensator by choosing $\widehat{\mathsf{F}}_n^{g,p}(a,b,-\infty) = \bar{\mathsf{F}}_n^{g,p}(a,b,-\infty) = 0$ while $\widehat{\mathsf{F}}_n^{g,p}(a,b,\infty) = \frac{1}{n}\sum_{i=1}^n g_{in}\varepsilon_i^p$ and $\bar{\mathsf{F}}_n^{g,p}(a,b,\infty) = \frac{1}{n}\sum_{i=1}^n g_{in}\mathsf{E}_{i-1}\varepsilon_i^p$.

We now define the empirical process, for $0 \le \psi \le 1$,

$$\mathbf{F}_n^{g,p}(a,b,c_\psi) = n^{1/2}\{\widehat{\mathsf{F}}_n^{g,p}(a,b,c_\psi) - \overline{\mathsf{F}}_n^{g,p}(a,b,c_\psi)\}. \tag{18}$$

We will show convergence that is uniform in a, b, c_ψ for the above process. This generalizes results in [22], which had no marks and no variation a in scale, in [20, Theorem 2.2.5], which had no marks, in [15, 17], which had marks, but no variation in quantile c and no variation a in scale respectively.

In the following, we first present the new result concerning variation in the scale a and the quantile c. Subsequently, we combine this with existing results concerning variation in b, c in order to get a result that is uniform in all three arguments a, b, c.

4.1 The Case of Estimated Scale and Known Regression Parameter

The main technical contribution of the paper is to analyze the empirical process in the case of estimated scale, but known regression parameter. Thus, we establish results for the empirical process that are uniform in a, c. Koul [20, Theorem 2.2.5] established a similar result for the case of unbounded weights g_{in} but no marks ε_i^p. His proof exploits that the function $1_{(\varepsilon_i \le \sigma c)}$ is monotone in c and bounded. These properties are not shared by $\varepsilon_i^p 1_{(\varepsilon_i \le \sigma c)}$, so we follow a different strategy for the proof that exploits the iterated martingale inequality from [18] reported as Lemma 3 in the Appendix 1.

We first present the uniformity result for the empirical process and then a uniform linearization result for the compensator. The proof involves a chaining argument. For this, we apply an iterated martingale inequality, see Lemma 3, to explore the tail behaviour of the maximum of a family of martingales.

Theorem 5 *Let $\mathscr{F}_i$ be an increasing sequence of σ-fields so ε_{i-1} and g_{in} are $\mathscr{F}_{i-1}$ measurable and ε_i is independent of $\mathscr{F}_{i-1}$. Let ε_i/σ have a continuous density* f. *Let p and η be given so $p \in \mathbf{N}_0$ and $0 < \eta \le 1/4$. Suppose*

(i) the density f *satisfies*

  (a) moments: $\int_{-\infty}^{\infty} |u|^{4p}\mathsf{f}(u)du < \infty$;
  (b) boundedness: $\sup_{u\in\mathbf{R}} |u|(1+|u|^{4p})\mathsf{f}(u) < \infty$;

(ii) the weights g_{in} satisfy $n^{-1}\mathsf{E}\sum_{i=1}^n |g_{in}|^4 = \mathrm{O}(1)$.

Let $c_\psi = \mathsf{F}^{-1}(\psi)$ for $0 \le \psi \le 1$. Then for any $B > 0$ and as $n \to \infty$

$$\sup_{0\le\psi\le 1}\ \sup_{|a|\le n^{1/4-\eta}B} |\mathbf{F}_n^{g,p}(a,0,c_\psi) - \mathbf{F}_n^{g,p}(0,0,c_\psi)| = \mathrm{o}_{\mathsf{P}}(1).$$

The second result provides a linearization of the compensator.

Theorem 6 *Let $\mathscr{F}_i$ be an increasing sequence of σ-fields so ε_{i-1} and g_{in} are $\mathscr{F}_{i-1}$ measurable and ε_i is independent of $\mathscr{F}_{i-1}$. Let ε_i/σ have a differentiable density* f. *Let p and η be given so $p \in \mathbf{N}_0$ and $0 < \eta \leq 1/4$. Suppose*

(i) the density f *satisfies*

(a) moments: $\int_{-\infty}^{\infty} |u|^p \mathsf{f}(u) du < \infty$;

(b) boundedness: $\sup_{u \in \mathbf{R}} u^2 |u^{p-1} \mathsf{f}(u) + u^p \dot{\mathsf{f}}(u)| < \infty$;

(ii) the weights g_{in} satisfy $n^{-1} \sum_{i=1}^{n} |g_{in}| = \mathrm{O}_{\mathsf{P}}(1)$.

Let $c_\psi = \mathsf{F}^{-1}(\psi)$ for $0 \leq \psi \leq 1$. Then for any $B > 0$ and as $n \to \infty$

$$\sup_{0 \leq \psi \leq 1} \sup_{|a| \leq n^{1/4-\eta} B} |n^{1/2} \{\overline{\mathsf{F}}_n^{g,p}(a, 0, c_\psi) - \overline{\mathsf{F}}_n^{g,p}(0, 0, c_\psi)\}$$
$$- \sigma^{p-1} c_\psi^p \mathsf{f}(c_\psi) n^{-1/2} \sum_{i=1}^{n} g_{in} n^{-1/2} a c_\psi | = \mathrm{O}_{\mathsf{P}}(n^{-2\eta}).$$

4.2 The Case of Estimated Scale and Regression Parameter

We now turn to the general one-sided empirical process with estimated scale and regression parameters. The case with known regression parameter was treated above while the case with known scale was treated in [18]. Through an argument reported in the appendix these results can be combined to prove the general result. For this we need the union of the various assumptions. This is listed below as Assumption 2. Note the density f is not necessarily symmetric in this section and Assumption 2 is weaker than Assumption 1.

Assumption 2 Let $\mathscr{F}_i$ be an increasing sequence of σ-fields so ε_{i-1}, x_i and g_{in} are $\mathscr{F}_{i-1}$ measurable and ε_i is independent of $\mathscr{F}_{i-1}$. Let ε_i/σ have a continuously differentiable density f which is positive on $\mathbf{R}$. Let p, η, κ be given so $p \in \mathbf{N}_0$ and $0 \leq \kappa < \eta \leq 1/4$. Choose $r \in \mathbf{N}_0$ so

$$2^{r-1} \geq 1 + (1/4 + \kappa - \eta)(1 + \dim x). \tag{19}$$

Suppose

(i) the density f satisfies

(a) moments: $\int_{-\infty}^{\infty} |u|^{2^r p} \mathsf{f}(u) du < \infty$;

(b) boundedness: $\sup_{u \in \mathbf{R}} [\{1 + |u|^{\max(4p+1, 2^r p - 1)}\} \mathsf{f}(u) + (1 + u^{2^r p + 2}) |\dot{\mathsf{f}}(u)|] < \infty$;

(c) smoothness: a $C_{\mathrm{H}} > 0$ exists so that for all $v > 0$

$$\frac{\sup_{u \geq v}(1+u^{2^r p})\mathsf{f}(u)}{\inf_{0 \leq u \leq v}(1+u^{2^r p})\mathsf{f}(u)} \leq C_{\mathrm{H}}, \qquad \frac{\sup_{u \leq -v}(1+u^{2^r p})\mathsf{f}(u)}{\inf_{-v \leq u \leq 0}(1+u^{2^r p})\mathsf{f}(u)} \leq C_{\mathrm{H}};$$

(ii) the regressors x_i satisfy $\max_{1 \leq i \leq n} |n^{1/2-\kappa} N' x_i| = \mathrm{O}_{\mathsf{P}}(1)$;
(iii) the weights g_{in} satisfy

(a) $n^{-1} \mathsf{E} \sum_{i=1}^{n} |g_{in}|^{2^r} (1 + |n^{1/2} N' x_i|) = \mathrm{O}(1)$;
(b) $n^{-1} \sum_{i=1}^{n} |g_{in}|(1 + |n^{1/2} N' x_i|^2) = \mathrm{O}_{\mathsf{P}}(1)$.

Remark 1 Assumption 1(*ia*, *iib*, *iic*) implies Assumption 2 with $r \geq 2$ satisfying (14) when g_{in} is either of 1, $n^{1/2} N' x_i$, $n N' x_i x_i' N$ and p is either of 0, 1, 2. Details are given in Lemma 4 in the appendix. □

We present two asymptotic results. The first theorem shows that the estimation error for the scale and regression parameter is negligible uniformly in the quantile.

Theorem 7 *Suppose Assumption 2 is satisfied. Let* $c_\psi = \mathsf{F}^{-1}(\psi)$ *for* $0 \leq \psi \leq 1$. *Then for any* $B > 0$ *and as* $n \to \infty$

$$\sup_{0 \leq \psi \leq 1} \sup_{|a|,|b| \leq n^{1/4-\eta} B} |\mathbf{F}_n^{g,p}(a, b, c_\psi) - \mathbf{F}_n^{g,p}(0, 0, c_\psi)| = \mathrm{o}_{\mathsf{P}}(1).$$

The proof has two parts. First, we keep a fixed and consider variation in b, c. This has been done in [17, Theorem 4.1]. Secondly, we keep b fixed and consider variation in a, c as done in Theorem 5.

The second result provides a linearization of the compensator.

Theorem 8 *Suppose Assumption* 2(*ia*, *ib*, *iiib*) *holds with* $r = 0$. *Let* $c_\psi = \mathsf{F}^{-1}(\psi)$ *for* $0 \leq \psi \leq 1$. *Then for any* $B > 0$ *and as* $n \to \infty$

$$\sup_{0 \leq \psi \leq 1} \sup_{|a|,|b| \leq n^{1/4-\eta} B} |n^{1/2}\{\overline{\mathsf{F}}_n^{g,p}(a, b, c_\psi) - \overline{\mathsf{F}}_n^{g,p}(0, 0, c_\psi)\}$$
$$-\sigma^{p-1} c_\psi^p \mathsf{f}(c_\psi) n^{-1/2} \sum_{i=1}^{n} g_{in}(n^{-1/2} a c_\psi + x_{in}' b)| = \mathrm{O}_{\mathsf{P}}(n^{-2\eta}).$$

Finally, the tightness of the empirical process $\mathbf{F}_n^{g,p}(0, 0, c_\psi)$ was shown in [17, Theorem 4.4], see tightness in [3].

4.3 A Result for the Two-Sided Empirical Process

The 1-step Huber-skip M-estimator involves indicators depending on the absolute value of the residuals. We therefore present some results for a class of two-sided weighted and marked empirical processes.

Define the weighted and marked absolute empirical distribution function

$$\widehat{\mathsf{G}}_n^{g,p}(a,b,c) = \frac{1}{n}\sum_{i=1}^n g_{in}\varepsilon_i^p 1_{(|\varepsilon_i - x_{in}'b| \le \sigma c + n^{-1/2}ac)}. \tag{20}$$

We suppose a so that $\sigma + n^{-1/2}a > 0$, in which case it suffices to consider $c \ge 0$. This restriction on a is satisfied when choosing a as $\widetilde{a} = n^{1/2}(\widetilde{\sigma} - \sigma)$ such that $\sigma + n^{-1/2}\widetilde{a} = \widetilde{\sigma} > 0$. Introduce the compensator of $\widehat{\mathsf{G}}_n^{g,p}(a,b,c)$

$$\overline{\mathsf{G}}_n^{g,p}(a,b,c) = \frac{1}{n}\sum_{i=1}^n g_{in}\mathsf{E}_{i-1}\varepsilon_i^p 1_{(|\varepsilon_i - x_{in}'b| \le \sigma c + n^{-1/2}ac)}. \tag{21}$$

Note $\overline{\mathsf{G}}_n^{1,0}(0,0,c) = \mathsf{G}(c) = \mathsf{P}(|\varepsilon_i| \le \sigma c)$. Then the absolute empirical process is

$$\mathbf{G}_n^{g,p}(a,b,c) = n^{1/2}\{\widehat{\mathsf{G}}_n^{g,p}(a,b,c) - \overline{\mathsf{G}}_n^{g,p}(a,b,c)\}. \tag{22}$$

We can now derive asymptotic theory for the absolute empirical process from Theorems 7 and 8. These results are presented under more restrictive Assumption 1, where the innovation distribution is symmetric, see Remark 1 and Lemma 4. In this section, we only consider g_{in} chosen as 1, $n^{1/2}N'x_i$, $nN'x_ix_i'N$ and p as 0, 1, 2.

Theorem 9 *Suppose Assumption 1(ia, iib, iic) holds. Let $c_\psi = \mathsf{G}^{-1}(\psi)$ for $0 \le \psi \le 1$. Then for all $B > 0$ and as $n \to \infty$*

$$\sup_{0\le\psi\le1}\ \sup_{|a|,|b|\le n^{1/4-\eta}B} |\mathbf{G}_n^{g,p}(a,b,c_\psi) - \mathbf{G}_n^{g,p}(0,0,c_\psi)| = \mathsf{o_P}(1).$$

Theorem 10 *Suppose Assumption 1(ia, iic) holds. Let $c_\psi = \mathsf{G}^{-1}(\psi)$ for $0 \le \psi \le 1$. Then for all $B > 0$ and as $n \to \infty$*

$$\sup_{0\le\psi\le1}\ \sup_{|a|,|b|\le n^{1/4-\eta}B} |n^{1/2}\{\overline{\mathsf{G}}_n^{g,p}(a,b,c_\psi) - \overline{\mathsf{G}}_n^{g,p}(0,0,c_\psi)\}$$
$$-2\sigma^{p-1}c_\psi^p\mathsf{f}(c_\psi)n^{-1/2}\sum_{i=1}^n g_{in}\{1_{(p\ even)}n^{-1/2}ac_\psi + 1_{(p\ odd)}x_{in}'b\}| = \mathsf{O_P}(n^{-2\eta}).$$

5 Discussion

This paper contributes to the asymptotic theory of iterated 1-step Huber-skip M-estimators. The results are derived under the null hypothesis that there are no outliers in the model. It is well known that the first-order asymptotic approximation is fragile in some small finite sample situations. Therefore, it would be of interest to carry out simulation studies to evaluate the finite sample performance of the results in this paper. Likewise it would be of interest to extend the result to situations where outliers are actually present in the data generating process. Scenario possibly contain single outliers, clusters of outliers, level shifts, symmetric or non-symmetric outliers. In such situations, we would analyze the potency, which is the retention rate for relevant outliers. Moreover, it would be possible to compare the potency of two distinct outlier detection algorithms with the same gauge.

Acknowledgements The second author is grateful to the Programme of Economic Modelling, which is part of Institute for New Economic Thinking at the Oxford Martin School. We thank Jana Jurečková and two anonymous referees for many constructive suggestions for improvement of the manuscript.

Appendix 1 A Metric on R and Some Inequalities

The asymptotic theory uses a chaining argument. This involves a partitioning of the quantile axis using a metric, which is presented first. Then follows some preliminary inequalities including an iterated exponential martingale inequality.

Define the function

$$J_{i,p}(x,y) = (\frac{\varepsilon_i}{\sigma})^p \{1_{(\varepsilon_i/\sigma \le y)} - 1_{(\varepsilon_i/\sigma \le x)}\}. \tag{23}$$

Our interest focus on $J_{i,p}(x,y)$ of order 2^r with $r \in \mathbf{N}$. Note that $u^{2^r p}$ is non-negative since $2^r p$ is even for $p \in \mathbf{N}_0$ and $r \in \mathbf{N}$. Introduce a positive and increasing function

$$\mathsf{H}_r(x) = \int_{-\infty}^{x} (1 + u^{2^r p}) \mathsf{f}(u) du. \tag{24}$$

The derivative of this function is $\dot{\mathsf{H}}_r(x) = (1 + x^{2^r p})\mathsf{f}(x)$. Then, denote the constant

$$H_r = \mathsf{H}_r(\infty) = \int_{-\infty}^{\infty} (1 + u^{2^r p}) \mathsf{f}(u) du, \tag{25}$$

which is finite by Assumption 2(*ia*). Selection of the specific $r \in \mathbf{N}$ will be more clear in proofs of the empirical process results. The intuition of $\mathsf{H}_r(x)$ is obtained through setting $p = 0$ so that $\mathsf{H}_r(x) = 2\mathsf{F}(x)$, $\dot{\mathsf{H}}_r(x) = 2\mathsf{f}(x)$ and $H_r = 2$. Therefore, $\mathsf{H}_r(x)$ is the generalization of the distribution $\mathsf{F}(x) \sim \varepsilon_i/\sigma$. For $x \le y$ and $0 \le s \le r$,

$$0 \le |\mathsf{E}\{J_{i,p}(x,y)^{2^s}\}| \le \mathsf{E}\{|J_{i,p}(x,y)|^{2^s}\} \le \mathsf{H}_r(y) - \mathsf{H}_r(x), \tag{26}$$

as $|u^p| < |u^q| + 1$ for $q \ge p \ge 0$. Let $|\mathsf{H}_r(x) - \mathsf{H}_r(y)|$ be the H_r-distance for $x, y \in \mathbf{R}$.

In the context of chaining, partition the range of $\mathsf{H}_r(c)$ into K intervals of equal size H_r/K. In other words, partition the support into K intervals by endpoints

$$-\infty = c_0 < c_1 < \cdots < c_{K-1} < c_K = \infty, \tag{27}$$

with $c_{-k} = c_0$ for $k \in \mathbf{N}$ so that for $1 \leq k \leq K$

$$\mathsf{H}_r(c_k) - \mathsf{H}_r(c_{k-1}) = \frac{H_r}{K}. \tag{28}$$

We first present two preliminary inequalities.

Lemma 1 *If* $|\tilde{c} - c| \leq |Ac + B|$ *and* $|A| \leq 1/2$*, then*

$$|c| \leq \frac{|\tilde{c}| + |B|}{1 - |A|}, \qquad (Ac + B)^2 \leq 16(A^2\tilde{c}^2 + B^2).$$

Proof (Lemma 1) First inequality. Since $|Ac + B| \leq |A||c| + |B|$, the assumption implies $c - |A||c| - |B| \leq \tilde{c} \leq c + |A||c| + |B|$. Suppose $c \geq 0$, then the lower inequality gives $c(1 - |A|) - |B| \leq \tilde{c}$ so that $c \leq (\tilde{c} + |B|)/(1 - |A|)$. Suppose $c < 0$, then the upper inequality gives $\tilde{c} \leq c(1 - |A|) + |B|$ so that $(\tilde{c} - |B|)/(1 - |A|) \leq c$. Combine to get $|c| \leq \max\{|(\tilde{c} + |B|)/(1 - |A|)|, |(\tilde{c} - |B|)/(1 - |A|)|\} \leq (|\tilde{c}| + |B|)/(1 - |A|)$.

Second inequality. The first inequality in the lemma, $(x + y)^2 \leq 2(x^2 + y^2)$ and $|A| \leq 1/2$ imply $c^2 \leq 8(\tilde{c}^2 + B^2)$ and $(Ac + B)^2 \leq 2(A^2c^2 + B^2)$. Combine to get $(Ac + B)^2 \leq 2(8A^2\tilde{c}^2 + 8A^2B^2 + B^2) \leq 16(A^2\tilde{c}^2 + B^2)$. □

The following lemma concerns the H_r-distance of multiplicative shifts.

Lemma 2 *Let* $r \in \mathbf{N}_0$. *Suppose* f *is a continuous density satisfying*

(a) moments: $\int_{-\infty}^{\infty} |u|^{2^r p}\mathsf{f}(u)du < \infty$*;*

(b) boundedness: $\sup_{c\in\mathbf{R}} |c|(1 + |c|^{2^r p})\mathsf{f}(c) < \infty$.

Let $c_\psi = \mathsf{F}^{-1}(\psi)$ *for* $0 \leq \psi \leq 1$. *Then, for any* $B > 0$*, there exists* $C > 0$ *so*

$$\sup_{0\leq\psi\leq1} \sup_{|a|\leq n^{1/4-\eta}B} \left|\mathsf{H}_r\left\{c_\psi\left(1 + n^{-1/2}a/\sigma\right)\right\} - \mathsf{H}_r(c_\psi)\right| \leq Cn^{-1/4-\eta}.$$

Proof (Lemma 2) Denote $\mathscr{H} = |\mathsf{H}_r\{c_\psi(1 + n^{-1/2}a/\sigma)\} - \mathsf{H}_r(c_\psi)|$. Apply the first order mean value theorem at the point c_ψ to get $\mathscr{H} = |\sigma^{-1}n^{-1/2}a||c_\psi||\dot{\mathsf{H}}_r(\tilde{c}_\psi)|$, where $|\tilde{c}_\psi - c_\psi| \leq |\sigma^{-1}n^{-1/2}ac_\psi|$ and $\dot{\mathsf{H}}_r(\tilde{c}_\psi) = (1 + \tilde{c}_\psi^{2^r p})\mathsf{f}(\tilde{c}_\psi)$.

There exists n_0, so for any $n > n_0$ we have $|\sigma^{-1}n^{-1/2}a| \leq 1/2$ uniformly in $|a| \leq n^{1/4-\eta}B$. First, for $n > n_0$, we apply the first inequality in Lemma 1 to obtain $|c_\psi| \leq |\tilde{c}_\psi|/(1 - |\sigma^{-1}n^{-1/2}a|) \leq 2|\tilde{c}_\psi|$. It follows

$$\mathscr{H} \leq \sigma^{-1}n^{-1/2}n^{1/4-\eta}B2|\tilde{c}_\psi||\dot{\mathsf{H}}_r(\tilde{c}_\psi)| \leq 2\sigma^{-1}B\sup_{c\in\mathbf{R}}|c||\dot{\mathsf{H}}_r(c)|n^{-1/4-\eta}.$$

Thus $\mathscr{H} \leq Cn^{-1/4-\eta}$ by condition (b) that $|c\dot{\mathsf{H}}_r(c)| = |c|(1+|c|^{2^r p})\mathsf{f}(c)$ is bounded uniformly in c.

Second, consider $n \leq n_0$. Note $\mathsf{H}_r(x) \leq \mathsf{H}_r(\infty) = H_r$ for any x so that the triangle inequality shows $\mathscr{H} \leq 2H_r$. With $0 < \eta \leq 1/4$, it follows

$$\mathscr{H} \leq 2H_r n^{1/4+\eta} n^{-1/4-\eta} \leq 2H_r n_0^{1/4+\eta} n^{-1/4-\eta} = Cn^{-1/4-\eta},$$

where $C = 2H_r n_0^{1/4+\eta}$ is finite since $H_r < \infty$ by condition (a). □

The chaining argument involves the tail behaviour of the maximum of a family of martingales which can be controlled using the following iterated martingale inequality taken from [17]. It builds on an exponential martingale inequality derived by Bercu and Touati [1, Theorem 2.1].

Lemma 3 ([17], Theorem 5.2) *For l so $1 \leq l \leq L$, let $z_{l,i}$ be $\mathscr{F}_i$ adapted satisfying* $\mathrm{E}z_{l,i}^{2^{\bar{r}}} < \infty$ *for some $\bar{r} \in \mathbf{N}$. Let $D_r = \max_{1\leq l\leq L}\sum_{i=1}^{n}\mathsf{E}_{i-1}z_{l,i}^{2^r}$ for $1 \leq r \leq \bar{r}$. Suppose, for some $\varsigma \geq 0$, $\lambda > 0$, that $L = \mathrm{O}(n^{\lambda})$ and $\mathrm{E}D_r = \mathrm{O}(n^{\varsigma})$ for $r \leq \bar{r}$. If $\upsilon > 0$ is chosen such that*

(i) $\varsigma < 2\upsilon$;
(ii) $\varsigma + \lambda < \upsilon 2^{\bar{r}}$;

then, for all $\kappa > 0$ and as $n \to \infty$

$$\lim_{n\to\infty} \mathsf{P}\left\{\max_{1\leq l\leq L} |\sum_{i=1}^{n}(z_{l,i} - \mathsf{E}_{i-1}z_{l,i})| > \kappa n^{\upsilon}\right\} = 0.$$

Appendix 2 Proofs of Empirical Process Results Concerning Scale

Here we prove the empirical process results concerning the variation in scale when the regression parameter is known. We use the distance function H_r with $r = 2$.

Proof (Theorem 5) Let $c_{\psi^\dagger} = c_\psi(1 + n^{-1/2}a/\sigma)$ so $\mathbf{F}_n^{g,p}(a, 0, c_\psi) = \mathbf{F}_n^{g,p}(0, 0, c_{\psi^\dagger})$. Note $c_{\psi^\dagger}$ can be greater or less than c_ψ, since a such that $|a| \leq n^{1/4-\eta}B$ and c_ψ can be either positive or negative. Assume $c_\psi < c_{\psi^\dagger}$ without loss of generality. Denote $R(c_\psi, c_{\psi^\dagger}) = \mathbf{F}_n^{g,p}(0, 0, c_{\psi^\dagger}) - \mathbf{F}_n^{g,p}(0, 0, c_\psi)$. The aim is to prove $\mathscr{R}_n = \mathrm{o}_{\mathsf{P}}(1)$ for $n \to \infty$ where $\mathscr{R}_n = \sup_{0\leq\psi\leq 1}\sup_{|a|\leq n^{1/4-\eta}B}|R(c_\psi, c_{\psi^\dagger})|$.

1. *Partition the support*. For $\delta, n > 0$ partition the range of quantiles c as laid out in (27) with $K = \mathrm{int}(H_r n^{1/2}/\delta)$ and $r = 2$ since $H_r < \infty$ by assumption (ia).
2. *Assign c_ψ and $c_{\psi^\dagger}$ to the partitioned support*. For each ψ and $\psi^\dagger$ there exist $k \leq k^\dagger$ and grid points so that $c_{k-1} < c_\psi \leq c_k$ and $c_{k^\dagger-1} < c_{\psi^\dagger} \leq c_{k^\dagger}$.

3. *Apply chaining.* Relate c_ψ to the nearest right grid point c_k and $c_{\psi^\dagger}$ to the nearest left grid point $c_{k^\dagger-1}$. Add and subtract $\mathbf{F}_n^{g,p}(0,0,c_k)$ and $\mathbf{F}_n^{g,p}(0,0,c_{k^\dagger-1})$ to $R(c_\psi, c_{\psi^\dagger})$. The triangle inequality gives

$$|R(c_\psi, c_{\psi^\dagger})| \leq |R(c_\psi, c_k)| + |R(c_k, c_{k^\dagger-1})| + |R(c_{k^\dagger-1}, c_{\psi^\dagger})|.$$

Note that if $c_\psi, c_{\psi^\dagger}$ are in the same interval, then $|R(c_k, c_{k^\dagger-1})| = |R(c_{k-1}, c_k)|$. If $c_\psi, c_{\psi^\dagger}$ are in the neighbouring intervals, then $|R(c_k, c_{k^\dagger-1})| = 0$. Apply chaining to obtain $\mathscr{R}_n \leq \mathscr{R}_{n,1} + \mathscr{R}_{n,2} + \mathscr{R}_{n,3} + \mathscr{R}_{n,4}$, where

$$\begin{aligned}
\mathscr{R}_{n,1} &= \max_{1\leq k<k^\dagger-1<K} |R(c_k, c_{k^\dagger-1})|,\\
\mathscr{R}_{n,2} &= \max_{1\leq k\leq K} |R(c_{k-1}, c_k)|,\\
\mathscr{R}_{n,3} &= \max_{1\leq k\leq K} \sup_{c_{k-1}<c_\psi\leq c_k} |R(c_\psi, c_k)|,\\
\mathscr{R}_{n,4} &= \max_{1\leq k^\dagger\leq K} \sup_{c_{k^\dagger-1}<c_{\psi^\dagger}\leq c_{k^\dagger}} |R(c_{k^\dagger-1}, c_{\psi^\dagger})|.
\end{aligned}$$

Thus, it suffices to show $\mathscr{R}_{n,j} = \mathsf{o}_\mathsf{P}(1)$ for $j = 1, 2, 3, 4$ as $n \to \infty$.

4. *The term $\mathscr{R}_{n,1}$ is $\mathsf{o}_\mathsf{P}(1)$.* Use Lemma 3 with $\upsilon = 1/2$. Let g_{in} have coordinates $g_{in}^* = \sigma^p g_{in}$. Recall the notation $J_{i,p}(x, y)$ in (23). Write the coordinates of $R(c_k, c_{k^\dagger-1})$ as $n^{-1/2}\sum_{i=1}^n(z_{l,i} - \mathsf{E}_{i-1}z_{l,i})$ with $z_{l,i} = g_{in}^* J_{i,p}(c_k, c_{k^\dagger-1})$, where l represents the indices $k, k^\dagger$ with $L \leq K^2$. Two conditions of Lemma 3 need to be verified.
The parameter λ. The set of indices l has the size $L = \mathrm{O}(n^\lambda)$ where $\lambda = 1$, since $L \leq K^2$ and $K = \mathrm{O}(n^{1/2})$.
The parameter ς. Consider $1 \leq s \leq r = 2$ (instead of $1 \leq r \leq \bar{r} = 2$). By construction of partition and assignment in steps 1, 2, then $c_\psi \leq c_k < c_{k^\dagger-1} < c_{\psi^\dagger}$. Thus,

$$\mathsf{E}_{i-1}J_{i,p}^{2^s}(c_k, c_{k^\dagger-1}) \leq \mathsf{H}_r(c_{k^\dagger-1}) - \mathsf{H}_r(c_k) \leq \mathsf{H}_r(c_{\psi^\dagger}) - \mathsf{H}_r(c_\psi) \leq Cn^{-1/4-\eta},$$

by Lemma 2 using assumption (i) for some finite $C > 0$. Since

$$D_s = \max_{1\leq l\leq L}\sum_{i=1}^n \mathsf{E}_{i-1}z_{l,i}^{2^s} = \max_{1\leq k<k^\dagger-1<K}\sum_{i=1}^n g_{in}^{*2^s}\mathsf{E}_{i-1}J_{i,p}^{2^s}(c_k, c_{k^\dagger-1}),$$

we then find $D_s \leq Cn^{-1/4-\eta}\sum_{i=1}^n g_{in}^{*2^s}$. Moreover, using assumption (ii) we find that $\mathsf{E}n^{-1}\sum_{i=1}^n g_{in}^{*2^s} = \mathrm{O}(1)$. Thus, with $\varsigma = 3/4 - \eta$, we have $\mathrm{E}D_s = \mathrm{O}(n^\varsigma)$.
Condition (i) is that $\varsigma < 2\upsilon$. This holds since $\eta > 0$ so $\varsigma = 3/4 - \eta < 1 = 2\upsilon$.
Condition (ii) is that $\varsigma + \lambda < \upsilon 2^r$ where $r = 2$. This is satisfied since $\eta > 0$ so $\varsigma + \lambda = 7/4 - \eta < 2 = \upsilon 2^r$.

5. *The term* $\mathscr{R}_{n,2}$ *is* $\mathsf{o_P}(1)$. Use Lemma 3 with $\upsilon = 1/2$ and $z_{l,i} = g^*_{in} J_{i,p}(c_{k-1}, c_k)$, where index $l = k$ has the size $L = K$. Two conditions of Lemma 3 need to be shown.
   The parameter λ. The size $L = \mathrm{O}(n^{\lambda})$ where $\lambda = 1/2$, since $L = K = \mathrm{O}(n^{1/2})$.
   The parameter ς. Consider $1 \le s \le r = 2$. The equality (28) shows

$$\mathsf{E}_{i-1} J^{2^s}_{i,p}(c_{k-1}, c_k) \le \mathsf{H}_r(c_k) - \mathsf{H}_r(c_{k-1}) = \frac{H_r}{K} = \mathrm{O}(n^{-1/2}).$$

   Then, we find

$$D_s = \max_{1\le l\le L} \sum_{i=1}^{n} \mathsf{E}_{i-1} z^{2^s}_{l,i} = \max_{1\le k\le K} \sum_{i=1}^{n} g^{*2^s}_{in} \mathsf{E}_{i-1} J^{2^s}_{i,p}(c_{k-1}, c_k) = \mathrm{O}(n^{-1/2}) \sum_{i=1}^{n} g^{*2^s}_{in}.$$

   It follows that $\mathrm{E} D_s = \mathrm{O}(n^{\varsigma})$ where $\varsigma = 1/2$ by assumption (*ii*).
   Condition (*i*) holds, since $\varsigma = 1/2 < 1 = 2\upsilon$.
   Condition (*ii*) holds, since $\varsigma + \lambda = 1 < 2 = \upsilon 2^r$.
6. *Decompose the term* $\mathscr{R}_{n,3}$. Apply the triangle and Jensen's inequality to obtain,

$$|R(c_\psi, c_k)| \le n^{-1/2} \sum_{i=1}^{n} |g^*_{in}|\{|J_{i,p}(c_\psi, c_k)| + \mathsf{E}_{i-1}|J_{i,p}(c_\psi, c_k)|\}.$$

   For $c_{k-1} < c_\psi \le c_k$ where $1 \le k \le K$, we have $|J_{i,p}(c_\psi, c_k)| \le |J_{i,p}(c_{k-1}, c_k)|$. Then,

$$\mathscr{R}_{n,3} \le \max_{1\le k\le K} n^{-1/2} \sum_{i=1}^{n} |g^*_{in}|\{|J_{i,p}(c_{k-1}, c_k)| + \mathsf{E}_{i-1}|J_{i,p}(c_{k-1}, c_k)|\}.$$

   Therefore, it can be argued that $\mathscr{R}_{n,3} \le \widetilde{\mathscr{R}}_{n,3} + 2\overline{\mathscr{R}}_{n,3}$, where

$$\widetilde{\mathscr{R}}_{n,3} = \max_{1\le k\le K} n^{-1/2} \sum_{i=1}^{n} |g^*_{in}|\{|J_{i,p}(c_{k-1}, c_k)| - \mathsf{E}_{i-1}|J_{i,p}(c_{k-1}, c_k)|\},$$

$$\overline{\mathscr{R}}_{n,3} = \max_{1\le k\le K} n^{-1/2} \sum_{i=1}^{n} |g^*_{in}|\mathsf{E}_{i-1}|J_{i,p}(c_{k-1}, c_k)|.$$

   Thus, it suffices to show $\widetilde{\mathscr{R}}_{n,3}$ and $\overline{\mathscr{R}}_{n,3}$ are $\mathsf{o_P}(1)$ as $n \to \infty$.
7. *The term* $\widetilde{\mathscr{R}}_{n,3}$ *is* $\mathsf{o_P}(1)$. Argue along the lines of step 5 to show $\widetilde{\mathscr{R}}_{n,3} = \mathsf{o_P}(1)$.

8. *Bounding the term* $\overline{\mathscr{R}}_{n,3}$. Use the equality (28) and $K = \mathrm{O}(H_r n^{1/2}/\delta)$ to get

$$\mathsf{E}_{i-1}|J_{i,p}(c_{k-1}, c_k)| \leq \mathsf{H}_r(c_k) - \mathsf{H}_r(c_{k-1}) = \frac{H_r}{K} = \mathrm{O}(n^{-1/2}\delta).$$

   We then find $\overline{\mathscr{R}}_{n,3} = \mathrm{O}(n^{-1/2}\delta)n^{-1/2}\sum_{i=1}^n |g_{in}^*| = \mathrm{O}_{\mathsf{P}}(\delta)$ by the Markov inequality and the assumption (*ii*) that $n^{-1}\sum_{i=1}^n \mathsf{E}|g_{in}^*|^4 = \mathrm{O}(1)$. Thus, choose δ sufficiently small so that $\overline{\mathscr{R}}_{n,3} = \mathrm{o}_{\mathsf{P}}(1)$.
9. *The term* $\mathscr{R}_{n,4}$ *is* $\mathrm{o}_{\mathsf{P}}(1)$. This is similar as to show $\mathscr{R}_{n,3} = \mathrm{o}_{\mathsf{P}}(1)$. Thus the same argument can be made through steps 6, 7, 8. □

Proof (Theorem 6) The term of interest is

$$D_n(a, c_\psi) = n^{1/2}\{\overline{\mathsf{F}}_n^{g,p}(a, 0, c_\psi) - \overline{\mathsf{F}}_n^{g,p}(0, 0, c_\psi)\}$$
$$-\sigma^{p-1}c_\psi^p \mathsf{f}(c_\psi)n^{-1/2}\sum_{i=1}^n g_{in}n^{-1/2}ac_\psi,$$

where $\overline{\mathsf{F}}_n^{g,p}$ is well-defined due to assumption (*ia*). Let $w_i^{a,c_\psi} = 1_{(\varepsilon_i \leq \sigma c_\psi + n^{-1/2}ac_\psi)} - 1_{(\varepsilon_i \leq \sigma c_\psi)}$ and $h_i(a, c_\psi) = n^{-1/2}ac_\psi/\sigma$ and denote $s(c) = c^p\mathsf{f}(c)$. Define $S_i(a, c_\psi) = \mathsf{E}_{i-1}\varepsilon_i^p w_i^{a,c_\psi} - \sigma^p h_i(a, c_\psi)s(c_\psi)$ so $D_n(a, c_\psi) = n^{-1/2}\sum_{i=1}^n g_{in}S_i(a, c_\psi)$. Write $S_i(a, c_\psi)$ as an integral and apply the second order Taylor expansion at c_ψ to get

$$S_i(a, c_\psi) = \sigma^p\left\{\int_{c_\psi}^{c_\psi + h_i(a,c_\psi)} s(u)du - h_i(a, c_\psi)s(c_\psi)\right\} = \sigma^p h_i^2(a, c_\psi)\dot{s}(\tilde{c}_\psi)/2,$$

where $|\tilde{c}_\psi - c_\psi| \leq |h_i(a, c_\psi)|$. There exists $n_0 > 0$ so for any $n > n_0$ we have $|\sigma^{-1}n^{-1/2}a| \leq 1/2$. We then apply the second inequality in Lemma 1 to obtain $h_i^2(a, c_\psi) \leq 16n^{-1}a^2\tilde{c}_\psi^2/\sigma^2$. Exploit the bound $|a| \leq n^{1/4-\eta}B$ to get

$$|S_i(a, c_\psi)| = \mathrm{O}(n^{-1/2-2\eta})\tilde{c}_\psi^2|\dot{s}(\tilde{c}_\psi)| = \mathrm{O}(n^{-1/2-2\eta})$$

uniformly in ψ, a, since $\tilde{c}_\psi^2|\dot{s}(\tilde{c}_\psi)| \leq \sup_{c\in\mathbf{R}} c^2|\dot{s}(c)| < \infty$ by assumption (*i*) noting that $\dot{s}(c) = c^{p-1}\mathsf{f}(c) + c^p\dot{\mathsf{f}}(c)$. Then the triangle inequality gives

$$|D_n(a, c_\psi)| \leq n^{-1/2}\sum_{i=1}^n |g_{in}||S_i(a, c_\psi)| = \mathrm{O}(n^{-2\eta})n^{-1}\sum_{i=1}^n |g_{in}|.$$

By assumption (*ii*), this term is of order $\mathrm{O}_{\mathsf{P}}(n^{-2\eta})$ uniformly in ψ, a. □

Appendix 3 Proofs of General Empirical Process Results

Proof (Theorem 7) The term of interest is $\mathscr{W} = \mathbf{F}_n^{g,p}(a, b, c_\psi) - \mathbf{F}_n^{g,p}(0, 0, c_\psi)$. Denote $c_{\psi^\dagger} = c_\psi(1 + n^{-1/2}a/\sigma)$. Notice that $\mathbf{F}_n^{g,p}(a, b, c_\psi) = \mathbf{F}_n^{g,p}(0, b, c_{\psi^\dagger})$ so that $\mathscr{W} = \mathbf{F}_n^{g,p}(0, b, c_{\psi^\dagger}) - \mathbf{F}_n^{g,p}(0, 0, c_\psi)$. Add and subtract $\mathbf{F}_n^{g,p}(a, 0, c_\psi) = \mathbf{F}_n^{g,p}(0, 0, c_{\psi^\dagger})$ and apply the triangle inequality to get

$$|\mathscr{W}| \le |\mathbf{F}_n^{g,p}(0, b, c_{\psi^\dagger}) - \mathbf{F}_n^{g,p}(0, 0, c_{\psi^\dagger})| + |\mathbf{F}_n^{g,p}(a, 0, c_\psi) - \mathbf{F}_n^{g,p}(0, 0, c_\psi)|.$$

Thus, the problem reduces to showing

$$\sup_{0\le\psi^\dagger\le 1} \sup_{|b|\le n^{1/4-\eta}B} |\mathbf{F}_n^{g,p}(0, b, c_{\psi^\dagger}) - \mathbf{F}_n^{g,p}(0, 0, c_{\psi^\dagger})| = \mathrm{o}_\mathsf{P}(1), \tag{29}$$

$$\sup_{0\le\psi\le 1} \sup_{|a|\le n^{1/4-\eta}B} |\mathbf{F}_n^{g,p}(a, 0, c_\psi) - \mathbf{F}_n^{g,p}(0, 0, c_\psi)| = \mathrm{o}_\mathsf{P}(1). \tag{30}$$

Then (29) is shown in [17, Theorem 4.1] by Assumption 2(i, ii, $iiia$) with $r \ge 2$ such that (14) holds. Further, (30) was considered in Theorem 5, which requires Assumption 2(ia, ib, iii) with $r = 2$. □

Proof (Theorem 8) We generalize the proof of Theorem 6. We note $\overline{\mathsf{F}}_n^{g,p}$ is well-defined due to Assumption 2(ia). The term of interest is

$$D_n(a, b, c_\psi) = n^{1/2}\{\overline{\mathsf{F}}_n^{g,p}(a, b, c_\psi) - \overline{\mathsf{F}}_n^{g,p}(0, 0, c_\psi)\}$$
$$-\sigma^{p-1}c_\psi^p \mathsf{f}(c_\psi)n^{-1/2}\sum_{i=1}^n g_{in}(n^{-1/2}ac_\psi + x_{in}'b).$$

Let $w_i^{a,b,c_\psi} = 1_{(\varepsilon_i\le\sigma c_\psi + n^{-1/2}ac_\psi + x_{in}'b)} - 1_{(\varepsilon_i\le\sigma c_\psi)}$, $h_i(a, b, c_\psi) = (n^{-1/2}ac_\psi + x_{in}'b)/\sigma$ and $s(c) = c^p\mathsf{f}(c)$. Define $S_i(a, b, c_\psi) = \mathsf{E}_{i-1}\varepsilon_i^p w_i^{a,b,c_\psi} - \sigma^p h_i(a, b, c_\psi)s(c_\psi)$ so that $D_n(a, b, c_\psi) = n^{-1/2}\sum_{i=1}^n g_{in}S_i(a, b, c_\psi)$. Write $S_i(a, b, c_\psi)$ as an integral

$$S_i(a, b, c_\psi) = \sigma^p\left\{\int_{c_\psi}^{c_\psi + h_i(a,b,c_\psi)} s(u)du - h_i(a, b, c_\psi)s(c_\psi)\right\}.$$

Second order Taylor expansion at c_ψ shows $S_i(a, b, c_\psi) = \sigma^p h_i^2(a, b, c_\psi)\dot{s}(\tilde{c}_\psi)/2$, where $|\tilde{c}_\psi - c_\psi| \le |h_i(a, b, c_\psi)|$. There exists $n_0 > 0$ so for any $n > n_0$ we have $|\sigma^{-1}n^{-1/2}a| \le 1/2$. We then apply the second inequality in Lemma 1 to obtain $h_i^2(a, b, c_\psi) \le 16\{n^{-1}a^2\tilde{c}_\psi^2 + (x_{in}'b)^2\}/\sigma^2$. Exploit bounds $|a|, |b| \le n^{1/4-\eta}B$ and the inequality $x^2 + y^2 \le (1 + x^2)(1 + y^2)$ to get

$$|S_i(a, b, c_\psi)| = \mathrm{O}(n^{-1/2-2\eta})(1 + |n^{1/2}x_{in}|^2)(1 + \tilde{c}_\psi^2)|\dot{s}(\tilde{c}_\psi)|.$$

Since $(1+\tilde{c}_\psi^2)|\dot{s}(\tilde{c}_\psi)| \leq \sup_{c\in\mathbf{R}}(1+c^2)|\dot{s}(c)| < \infty$ by Assumption 2(*ib*) with $r=0$, we have $|S_i(a,b,c_\psi)| = \mathrm{O}(n^{-1/2-2\eta})(1+|n^{1/2}x_{in}|^2)$ uniformly in ψ, a, b. Then the triangle inequality gives

$$|D_n(a,b,c_\psi)| \leq n^{-1/2}\sum_{i=1}^n |g_{in}||S_i(a,b,c_\psi)| = \mathrm{O}(n^{-2\eta})n^{-1}\sum_{i=1}^n |g_{in}|(1+|n^{1/2}x_{in}|^2).$$

By Assumption 2(*iiib*), this term is of order $\mathrm{O}_\mathsf{P}(n^{-2\eta})$ uniformly in ψ, a, b. □

The absolute empirical process results are given under more restrictive Assumption 1, so the next lemma concerns the relationship between Assumptions 1 and 2.

Lemma 4 *Suppose g_{in} is either of* 1, $n^{1/2}N'x_i$, $nN'x_ix_i'N$ *and p is either of* 0, 1, 2. *Then Assumption 1*(*ia*, *iib*, *iic*) *implies Assumption 2 with $r \geq 2$ satisfying (14).*

Proof (Lemma 4) Assumption 1(*ia*) shows Assumption 2(*ia*, *ic*), while Assumption 2(*ib*) further needs continuous differentiability of f, see discussion in [17, Remark 4.1(c)]. Assumption 1(*iib*) is the same as Assumption 2(*ii*). Assumption 1(*iic*) implies Assumption 2(*iiia*) and (*iiic*) by Markov inequality. □

Proof (Theorem 9) The term of interest is $\mathscr{G} = \mathbf{G}_n^{g,p}(a,b,c_\psi) - \mathbf{G}_n^{g,p}(0,0,c_\psi)$. Our focus is on the absolute quantile $c_\psi = \mathsf{G}^{-1}(\psi) > 0$ rather than the one-sided quantile $c_{\psi^*} = \mathsf{F}^{-1}(\psi^*) \in \mathbf{R}$. Note $|\varepsilon_i|/\sigma \sim \mathsf{G}$ and $\varepsilon_i/\sigma \sim \mathsf{F}$. Since

$$1_{(|\varepsilon_i - x_{in}'b| \leq \sigma c + n^{-1/2}ac)} = 1_{(\varepsilon_i \leq \sigma c + n^{-1/2}ac + x_{in}'b)} - 1_{(\varepsilon_i \leq -\sigma c - n^{-1/2}ac + x_{in}'b)}$$

and by (18) and (22), we have $\mathbf{G}_n^{g,p}(a,b,c) = \mathbf{F}_n^{g,p}(a,b,c) - \lim_{c^\dagger\downarrow c}\mathbf{F}_n^{g,p}(a,b,-c^\dagger)$ for any $c > 0$. By this and the triangle inequality, then for any $c_\psi = \mathsf{G}^{-1}(\psi) > 0$,

$$|\mathscr{G}| \leq |\mathbf{F}_n^{g,p}(a,b,c_\psi) - \mathbf{F}_n^{g,p}(0,0,c_\psi)| + \lim_{c_\psi^\dagger \downarrow c_\psi} |\mathbf{F}_n^{g,p}(a,b,-c_\psi^\dagger) - \mathbf{F}_n^{g,p}(0,0,-c_\psi^\dagger)|.$$

These vanish uniformly in ψ, a, b by Theorem 7 using Assumption 2 with $r \geq 2$ such that (14) holds. Lemma 4 shows that Assumption 1(*ia*, *iib*, *iic*) suffices. □

Proof (Theorem 10) Argue as in the proof of Theorem 9 but using Theorem 8 instead of Theorem 7. Due to the symmetry of f, the correction term is then

$$\sigma^{p-1}c_\psi^p\mathsf{f}(c_\psi)n^{-1/2}\sum_{i=1}^n g_{in}[\{1+(-1)^p\}n^{-1/2}ac_\psi + \{1-(-1)^p\}x_{in}'b].$$

This reduces as desired. □

Appendix 4 Proofs of the Main Results

We first present an axillary result for asymptotic expansions of product moments. Then, the tightness and fixed point result are shown for the iterated estimators. At last, we provide the proof of the Poisson exceedence theory for the gauge.

The 1-step Huber-skip M-estimators are least squares estimators for selected observations. The following result describes the asymptotic behaviour of the corresponding product moments. For this purpose introduce the indicators

$$v_i^{a,b,c} = 1_{(|\varepsilon_i - x_{in}' b| \le \sigma c + n^{-1/2} ac)}. \tag{31}$$

Lemma 5 *Suppose Assumption 1(ia, ii) holds. Then we have expansions*

$$n^{-1/2}\sum_{i=1}^n v_i^{a,b,c} = n^{-1/2}\sum_{i=1}^n 1_{(|\varepsilon_i| \le \sigma c)} + 2\mathsf{f}(c)\frac{ac}{\sigma} + R_v(a,b,c),$$

$$n^{-1/2}\sum_{i=1}^n \varepsilon_i^2 v_i^{a,b,c} = n^{-1/2}\sum_{i=1}^n \varepsilon_i^2 1_{(|\varepsilon_i| \le \sigma c)} + 2\sigma^2 c^2 \mathsf{f}(c)\frac{ac}{\sigma} + R_{v\varepsilon\varepsilon}(a,b,c),$$

$$\sum_{i=1}^n N' x_i \varepsilon_i v_i^{a,b,c} = \sum_{i=1}^n N' x_i \varepsilon_i 1_{(|\varepsilon_i| \le \sigma c)} + 2c\mathsf{f}(c)\Sigma b + R_{vx\varepsilon}(a,b,c),$$

$$n^{1/2}\sum_{i=1}^n N' x_i x_i' N v_i^{a,b,c} = n^{1/2}\sum_{i=1}^n N' x_i x_i' N 1_{(|\varepsilon_i| \le \sigma c)} + 2\mathsf{f}(c)\Sigma\frac{ac}{\sigma} + R_{vxx}(a,b,c).$$

Let $R(a,b,c) = |R_v(a,b,c)| + |R_{v\varepsilon\varepsilon}(a,b,c)| + |R_{vx\varepsilon}(a,b,c)| + |R_{vxx}(a,b,c)|$. *Then for any* $B > 0$ *and as* $n \to \infty$

$$\sup_{0<c<\infty}\ \sup_{|a|,|b| \le n^{1/4-\eta}B} |R(a,b,c)| = \mathrm{o}_{\mathsf{P}}(1).$$

Remark 2 The first and fourth item in Lemma 5 adjusted by $n^{-1/2}$ have expansions

$$n^{-1}\sum_{i=1}^n v_i^{a,b,c} = \psi + R_v'(a,b,c), \qquad \sum_{i=1}^n N' x_i x_i' N v_i^{a,b,c} = \psi\Sigma + R_{vxx}'(a,b,c),$$

where for any $B > 0$ and as $n \to \infty$

$$\sup_{0<c<\infty}\ \sup_{|a|,|b| \le n^{1/4-\eta}B} |R_v'(a,b,c)| + |R_{vxx}'(a,b,c)| = \mathrm{o}_{\mathsf{P}}(1).$$

Indeed, for the first expansion, we apply the law of large numbers to obtain $n^{-1}\sum_{i=1}^{n} 1_{(|\varepsilon_i|\le\sigma c)} = \psi + \mathsf{o}_{\mathsf{P}}(1)$, while $\sup_{c\in\mathbf{R}}|c|\mathsf{f}(c) < \infty$ by Assumption 1(*ia*) and $n^{-1/2}a$ vanishes. For the second expansion, decompose

$$\sum_{i=1}^{n} N'x_i x_i' N 1_{(|\varepsilon_i|\le\sigma c)} = \sum_{i=1}^{n} N'x_i x_i' N\{1_{(|\varepsilon_i|\le\sigma c)} - \psi\} + \sum_{i=1}^{n} N'x_i x_i' N\psi.$$

The first item vanishes by the Chebyshev inequality and Assumption 1(*iia*, *iic*), while the second converges to $\psi\Sigma$. □

Proof (Lemma 5) The general class of empirical processes is

$$\mathscr{M}_n = n^{-1/2}\sum_{i=1}^{n} g_{in}\varepsilon_i^p v_i^{a,b,c}, \quad v_i^{a,b,c} = 1_{(|\varepsilon_i - x_{in}'b|\le\sigma c + n^{-1/2}ac)}.$$

1. *Decompose* $\mathscr{M}_n$. Write $\mathscr{M}_n = \mathscr{M}_{n,1} + \mathscr{M}_{n,2} + \mathscr{M}_{n,3}$, where

$$\mathscr{M}_{n,1} = n^{-1/2}\sum_{i=1}^{n} g_{in}\varepsilon_i^p 1_{(|\varepsilon_i|\le\sigma c)}, \quad \mathscr{M}_{n,2} = n^{-1/2}\sum_{i=1}^{n} g_{in}\mathsf{E}_{i-1}\varepsilon_i^p\{v_i^{a,b,c} - 1_{(|\varepsilon_i|\le\sigma c)}\},$$

$$\mathscr{M}_{n,3} = n^{-1/2}\sum_{i=1}^{n} g_{in}\varepsilon_i^p\{v_i^{a,b,c} - 1_{(|\varepsilon_i|\le\sigma c)}\} - n^{-1/2}\sum_{i=1}^{n} g_{in}\mathsf{E}_{i-1}\varepsilon_i^p\{v_i^{a,b,c} - 1_{(|\varepsilon_i|\le\sigma c)}\}.$$

Therefore, the first term in stochastic expansion is $\mathscr{M}_{n,1}$. We will linearize $\mathscr{M}_{n,2}$ to obtain the second term, and argue that $\mathscr{M}_{n,3}$ is small in probability.

2. *Linearize* $\mathscr{M}_{n,2}$. Note $\mathscr{M}_{n,2} = n^{1/2}\{\overline{\mathsf{G}}_n^{g,p}(a,b,c) - \overline{\mathsf{G}}_n^{g,p}(0,0,c)\}$, see (21). Theorem 10 by Assumption 1(*ia*, *iic*) shows $\mathscr{M}_{n,2} = \overline{\mathscr{M}}_{n,2} + \mathsf{O}_{\mathsf{P}}(n^{-2\eta})$, where

$$\overline{\mathscr{M}}_{n,2} = 2\sigma^{p-1}c^p\mathsf{f}(c)n^{-1/2}\sum_{i=1}^{n} g_{in}\{1_{(p\ even)}n^{-1/2}ac + 1_{(p\ odd)}x_{in}'b\}.$$

This reduces as desired by Assumption 1(*iia*). Note $0 < \eta \le 1/4$. Thus, we have $\mathscr{M}_{n,2} = \overline{\mathscr{M}}_{n,2} + \mathsf{o}_{\mathsf{P}}(1)$ uniformly in $0 < c < \infty$ and $|a|, |b| \le n^{1/4-\eta}B$.

3. *Bounding* $\mathscr{M}_{n,3}$. Note $\mathscr{M}_{n,3} = \mathbf{G}_n^{g,p}(a,b,c) - \mathbf{G}_n^{g,p}(0,0,c)$, see (22). Due to Assumption 1(*ia*, *iib*, *iic*), Theorem 9 shows $\mathscr{M}_{n,3} = \mathsf{o}_{\mathsf{P}}(1)$ uniformly in a, b, c. □

Proof (Theorem 1) The $m+1$ step estimators for β, σ^2 are defined in (10), (11). These are least squares estimators for the non-outlying observations and satisfy

$$N^{-1}(\widehat{\beta}_c^{(m+1)} - \beta) = \left(\sum_{i=1}^n N' x_i x_i' N v_{i,c}^{(m)}\right)^{-1} \left(\sum_{i=1}^n N' x_i \varepsilon_i v_{i,c}^{(m)}\right), \tag{32}$$

$$n^{1/2}\{(\widehat{\sigma}_c^{(m+1)})^2 - \sigma^2\} = \varsigma_c^{-2}\left(n^{-1}\sum_{i=1}^n v_{i,c}^{(m)}\right)^{-1} n^{-1/2}\left\{\sum_{i=1}^n (\varepsilon_i^2 - \varsigma_c^2\sigma^2) v_{i,c}^{(m)} \right. \tag{33}$$
$$\left. - \left(\sum_{i=1}^n \varepsilon_i x_i' N v_{i,c}^{(m)}\right)\left(\sum_{i=1}^n N' x_i x_i' N v_{i,c}^{(m)}\right)^{-1}\left(\sum_{i=1}^n N' x_i \varepsilon_i v_{i,c}^{(m)}\right)\right\}.$$

We express the weight $v_{i,c}^{(m)}$ in (9) as

$$v_{i,c}^{(m)} = 1_{(|y_i - x_i'\widehat{\beta}_c^{(m)}| \le \widehat{\sigma}_c^{(m)} c)} = 1_{(|\varepsilon_i - x_{in}'\widehat{b}_c^{(m)}| \le \sigma c + n^{-1/2}\widehat{a}_c^{(m)} c)} = v_i^{\widehat{a}_c^{(m)}, \widehat{b}_c^{(m)}, c},$$

where $\widehat{b}_c^{(m)} = N^{-1}(\widehat{\beta}_c^{(m)} - \beta)$ and $\widehat{a}_c^{(m)} = n^{1/2}(\widehat{\sigma}_c^{(m)} - \sigma)$ are the m step estimation errors for β and σ.

Since $|\widehat{b}_c^{(m)}| + |\widehat{a}_c^{(m)}| = \mathsf{O_P}(1)$ and by Assumption 1(ia, ii), then Lemma 5 and Remark 2 with $\kappa = 0$, $\eta = 1/4$ show asymptotic expansions for product moments. Substitute these expansions into (32), (33) to first get

$$\widehat{b}_c^{(m+1)} = \frac{2c\mathsf{f}(c)}{\psi}\widehat{b}_c^{(m)} + (\psi\Sigma)^{-1}\sum_{i=1}^n N' x_i \varepsilon_i 1_{(|\varepsilon_i| \le \sigma c)} + R_\beta(\widehat{a}_c^{(m)}, \widehat{b}_c^{(m)}, c),$$

where the remainder $R_\beta(a, b, c)$ vanishes uniformly in $c_0 \le c < \infty$ and $|a|, |b| \le B$. A key to this is that c is bounded away from zero and that Σ is positive definite by Assumption 1(iia) so that the denominator ψ, $\psi\Sigma$ is bounded away from zero.

Secondly, we get an expression for $\widehat{\sigma}_c^{(m+1)}$. By Taylor expansion, first note that

$$n^{1/2}(\widehat{\sigma}_c^{(m+1)} - \sigma) = \frac{1}{2\sigma}n^{1/2}\{(\widehat{\sigma}_c^{(m+1)})^2 - \sigma^2\} + n^{-1/2}\mathrm{O}[n\{(\widehat{\sigma}_c^{(m+1)})^2 - \sigma^2\}^2].$$

Then apply arguments as above to get

$$\widehat{a}_c^{(m+1)} = \frac{c(c^2 - \varsigma_c^2)\mathsf{f}(c)}{\tau_2^c}\widehat{a}_c^{(m)} + \frac{1}{2\sigma\tau_2^c}n^{-1/2}\sum_{i=1}^n (\varepsilon_i^2 - \varsigma_c^2\sigma^2)1_{(|\varepsilon_i| \le \sigma c)}$$
$$+ R_\sigma(\widehat{a}_c^{(m)}, \widehat{b}_c^{(m)}, c),$$

where the remainder $R_\sigma(a, b, c)$ also vanishes uniformly. □

To prove the tightness and fixed point result, let $|\cdot|$ refer to the usual Euclidean vector norm, while $\|M\| = \max\{\text{eigen}(M'M)\}^{1/2}$ is the spectral norm for any matrix M. Note that the norms are compatible so that $|Mx| \le \|M\||x|$ for any vector x.

Proof (Theorem 2) Due to Assumption 1(ia, ii), Theorem 1 shows

$$\widehat{u}_c^{(m+1)} = \Gamma_c \widehat{u}_c^{(m)} + K_c + R_u(\widehat{u}_c^{(m)}, c), \tag{34}$$

where the remainder term satisfies $\sup_{c_0 \le c < \infty} \sup_{|u| \le B} |R_u(u, c)| = \mathrm{o}_{\mathsf{P}}(1)$ and

$$\widehat{u}_c^{(m)} = \begin{pmatrix} \widehat{b}_c^{(m)} \\ \widehat{a}_c^{(m)} \end{pmatrix} = \left\{ \begin{matrix} N^{-1}(\widehat{\beta}_c^{(m)} - \beta) \\ n^{1/2}(\widehat{\sigma}_c^{(m)} - \sigma) \end{matrix} \right\}, \quad \Gamma_c = \left\{ \begin{matrix} \frac{2c\mathsf{f}(c)}{\psi} I_{\dim x} & 0 \\ 0 & \frac{c(c^2 - \varsigma_c^2)\mathsf{f}(c)}{\tau_2^c} \end{matrix} \right\}, \tag{35}$$

$$K_c = \left\{ \begin{matrix} (\psi \Sigma)^{-1} & 0 \\ 0 & (2\sigma \tau_2^c)^{-1} \end{matrix} \right\} \sum_{i=1}^{n} \left\{ \begin{matrix} N' x_i \varepsilon_i \\ n^{-1/2}(\varepsilon_i^2 - \varsigma_c^2 \sigma^2) \end{matrix} \right\} 1_{(|\varepsilon_i| \le \sigma c)}. \tag{36}$$

Apply the difference Eq. (34) recursively to obtain the representation

$$\widehat{u}_c^{(m+1)} = \Gamma_c^{m+1} \widehat{u}_c^{(0)} + \sum_{l=0}^{m} \Gamma_c^l \{K_c + R_u(\widehat{u}_c^{(m-l)}, c)\}. \tag{37}$$

Use the triangle inequality and $|Mx| \le \|M\||x|$ to get

$$|\widehat{u}_c^{(m+1)}| \le \|\Gamma_c^{m+1}\||\widehat{u}_c^{(0)}| + \{|K_c| + \max_{0 \le l \le m} |R_u(\widehat{u}_c^{(l)}, c)|\} \sum_{l=0}^{m} \|\Gamma_c^l\|.$$

Assumption 1(ia) shows $\sup_{c_0 \le c < \infty} \max\{|2c\mathsf{f}(c)/\psi|, |c(c^2 - \varsigma_c^2)\mathsf{f}(c)/\tau_2^c|\} < 1$, see [16, Theorem 3.5], so $\sup_{c_0 \le c < \infty} \|\Gamma_c\| < 1$. Gelfand's formula in [24, Theorem 3.4] gives $\lim_{m \to \infty} \|M^m\|^{1/m} = \max |\mathrm{eigen}(M)|$. Therefore for some ω satisfying that $\sup_{c_0 \le c < \infty} \|\Gamma_c\| < \omega < 1$ there exists $m_0 > 0$ so for all $m > m_0$

$$\sup_{c_0 \le c < \infty} \|\Gamma_c^m\| < \omega^m < 1. \tag{38}$$

Also note $(I_{\dim x + 1} - \Gamma_c)^{-1} = \sum_{l=0}^{\infty} \Gamma_c$. This in turn implies for some $1 < B_0 < \infty$

$$\sup_{0 \le m < \infty} \sup_{c_0 \le c < \infty} \|\Gamma_c^m\| < B_0, \quad \sup_{c_0 \le c < \infty} \|(I_{\dim x + 1} - \Gamma_c)^{-1}\| \le \sum_{l=0}^{\infty} \sup_{c_0 \le c < \infty} \|\Gamma_c^l\| < B_0. \tag{39}$$

Therefore, (39) shows for all $m \in [0, \infty)$

$$|\widehat{u}_c^{(m+1)}| < B_0\{|\widehat{u}_c^{(0)}| + |K_c| + \max_{0 \le l \le m} |R_u(\widehat{u}_c^{(l)}, c)|\}. \tag{40}$$

For any $c \in [c_0, \infty)$, Assumption 1(iii) with $\eta = 1/4$ guarantees tightness of $\widehat{u}_c^{(0)}$, while the kernel K_c is tight by [17, Theorem 4.4] using Assumption 1(ia, iib, iic). Thus, for all $\varepsilon, \delta > 0$ there exist $n_0, U_0 > 0$ so that the set

$$\mathscr{A}_n = \{B_0 \sup_{c_0 \le c < \infty} (|\widehat{u}_c^{(0)}| + |K_c|) \le U_0/3, B_0 \sup_{c_0 \le c < \infty} \sup_{|u| \le U_0} |R_u(u, c)| < \delta/2\} \tag{41}$$

has probability larger than $1 - \varepsilon$ for all $n > n_0$.

Mathematical induction over m is used to show $\sup_{0\leq m<\infty}\sup_{c_0\leq c<\infty}|\widehat{u}_c^{(m)}|\leq U_0$ on the set $\mathscr{A}_n$. For $m=0$ as induction starts, $\sup_{c_0\leq c<\infty}|\widehat{u}_c^{(0)}|\leq B_0^{-1}U_0/3<U_0$ holds since $B_0>1$. The induction assumption is that $\sup_{0\leq l\leq m}\sup_{c_0\leq c<\infty}|\widehat{u}_c^{(l)}|\leq U_0$. This implies $B_0\max_{0\leq l\leq m}|R_u(\widehat{u}_c^{(l)},c)|<\delta/2$, and then the bound in (40) becomes $\sup_{c_0\leq c<\infty}|\widehat{u}_c^{(m+1)}|<2U_0/3+\delta/2<U_0$ so $\sup_{0\leq l\leq m+1}\sup_{c_0\leq c<\infty}|\widehat{u}_c^{(l)}|\leq U_0$. □

Proof (Theorem 3) Due to Assumption 1(*ia*, *ii*, *iii*), Theorem 1 provides the recursive Eq. (34). Then, Theorem 2 shows $\sup_{0\leq m<\infty}\sup_{c_0\leq c<\infty}|\widehat{u}_c^{(m)}|=\mathrm{O}_{\mathsf{P}}(1)$, so the remainder term in (34) is $\mathrm{o}_{\mathsf{P}}(1)$. Thus, for $m,n\to\infty$ the fixed point should satisfy $\widehat{u}_c^*=\Gamma_c\widehat{u}_c^*+K_c$ so that

$$\widehat{u}_c^*=(I_{\dim x+1}-\Gamma_c)^{-1}K_c. \tag{42}$$

Substitute (35), (36) of $\widehat{u}_c^*$, Γ_c and K_c into (42) to obtain

$$\begin{Bmatrix} N^{-1}(\widehat{\beta}_c^*-\beta) \\ n^{1/2}(\widehat{\sigma}_c^*-\sigma) \end{Bmatrix}=\begin{bmatrix} \frac{1}{\psi-2c\mathsf{f}(c)}\Sigma^{-1}\sum_{i=1}^n N'x_i\varepsilon_i 1_{(|\varepsilon_i|\leq\sigma c)} \\ \frac{1}{2\sigma\{\tau_2^c-c(c^2-\varsigma_c^2)\mathsf{f}(c)\}}n^{-1/2}\sum_{i=1}^n(\varepsilon_i^2-\varsigma_c^2\sigma^2)1_{(|\varepsilon_i|\leq\sigma c)} \end{bmatrix}.$$

Replace (37) and (42) into the deviation $\widehat{\Delta}_c^{(m+1)}=\widehat{u}_c^{(m+1)}-\widehat{u}_c^*$, and then apply $\sum_{l=0}^m\Gamma_c^l=(I_{\dim x+1}-\Gamma_c^{m+1})(I_{\dim x+1}-\Gamma_c)^{-1}$ to attain

$$\widehat{\Delta}_c^{(m+1)}=\Gamma_c^{m+1}\{\widehat{u}_c^{(0)}-(I_{\dim x+1}-\Gamma_c)^{-1}K_c\}+\sum_{l=0}^m\Gamma_c^l R_u(\widehat{u}_c^{(m-l)},c).$$

To bound $\widehat{\Delta}_c^{(m+1)}$, use the triangle inequality and $|Mx|\leq\|M\||x|$ to get

$$|\widehat{\Delta}_c^{(m+1)}|\leq\|\Gamma_c^{m+1}\|\{|\widehat{u}_c^{(0)}|+\|(I_{\dim x+1}-\Gamma_c)^{-1}\||K_c|\}+\max_{0\leq l\leq m}|R_u(\widehat{u}_c^{(l)},c)|\sum_{l=0}^m\|\Gamma_c^l\|.$$

By Assumption 1(*ia*) and Gelfand's formula, (38) and the second inequality in (39) imply for $m>m_0$

$$|\widehat{\Delta}_c^{(m+1)}|<\omega^{m+1}(|\widehat{u}_c^{(0)}|+B_0|K_c|)+B_0\max_{0\leq l\leq m}|R_u(\widehat{u}_c^{(l)},c)|.$$

On the set $\mathscr{A}_n$ as in (41), since $\sup_{0\leq m<\infty}\sup_{c_0\leq c<\infty}|\widehat{u}_c^{(m)}|\leq U_0$ by Theorem 2, we then have $\sup_{c_0\leq c<\infty}|\widehat{\Delta}_c^{(m+1)}|<\omega^{m+1}(B_0^{-1}U_0/3+U_0/3)+\delta/2<\omega^{m+1}U_0+\delta/2$. As $0<\omega<1$, ω^{m+1} declines exponentially so m_0 can be chosen sufficiently large that for all $m>m_0$ then $\omega^{m+1}U_0<\delta/2$. Thus $\mathsf{P}(\sup_{c_0\leq c<\infty}|\widehat{\Delta}_c^{(m+1)}|<\delta)>1-\varepsilon$ for all $m>m_0$ and $n>n_0$. □

Proof (Theorem 4) Assumption 1(*ia*) implies $\mathsf{E}|\varepsilon_i/\sigma|^l<\infty$ for some $l>4$. Apply (15) and the Chebyshev inequality to get $\lambda/n=\mathsf{P}(|\varepsilon_i|>\sigma c_n)\leq\mathsf{E}|\varepsilon_i/\sigma|^l c_n^{-l}$. Thus $c_n\leq(\mathsf{E}|\varepsilon_i/\sigma|^l)^{1/l}\lambda^{-1/l}n^{1/l}$ so that the divergence rate of c_n is $\mathrm{O}(n^{1/l})=\mathrm{o}(n^{1/4})$.

1. *A bound on the sample space*. By Assumption 1(ia, ii, iii) with $\eta = 1/4$, Theorems 2 and 3 show that $\widehat{\beta}_{c_n}^{(m)}$, $(\widehat{\sigma}_{c_n}^{(m)})^2$ are tight. Assumption 1(iib) gives $\max_{1\le i\le n}|x_{in}| = \mathsf{O}_\mathsf{P}(n^{\kappa-1/2}) = \mathsf{o}_\mathsf{P}(n^{-1/4})$ for some $0 \le \kappa < 1/4$. Thus, for all $\varepsilon > 0$ there exists a large constant A_0 so that the set

$$\mathscr{B}_n = \{\sup_{0\le m<\infty}(|\widehat{b}_{c_n}^{(m)}| + |\widehat{a}_{c_n}^{(m)}|) + n^{1/4}\max_{1\le i\le n}|x_{in}| \le A_0\}$$

has the probability larger than $1-\varepsilon$ for all n. Note that $\widehat{b}_{c_n}^{(m)} = N^{-1}(\widehat{\beta}_{c_n}^{(m)} - \beta)$ and $\widehat{a}_{c_n}^{(m)} = n^{1/2}(\widehat{\sigma}_{c_n}^{(m)} - \sigma)$.

2. *Bound the indicator*. Define the random quantity,

$$s_{i,c_n}^{(m)} = \widehat{\sigma}_{c_n}^{(m)}c_n - y_i + x_i'\widehat{\beta}_{c_n}^{(m)} + \varepsilon_i = \sigma c_n + n^{-1/2}\widehat{a}_{c_n}^{(m)}c_n + x_{in}'\widehat{b}_{c_n}^{(m)}.$$

On the set $\mathscr{B}_n$ and as $c_n = \mathrm{o}(n^{1/4})$, we have for some $A_1 > 0$

$$s_{i,c_n}^{(m)} \le \sigma c_n + n^{-1/2}A_0c_n + n^{-1/4}A_0^2 \le \sigma(c_n + n^{-1/4}A_1),$$
$$s_{i,c_n}^{(m)} \ge \sigma c_n - n^{-1/2}A_0c_n - n^{-1/4}A_0^2 \ge \sigma(c_n - n^{-1/4}A_1).$$

Since the sets $(y_i - x_i'\widehat{\beta}_{c_n}^{(m)} > \widehat{\sigma}_{c_n}^{(m)}c_n)$ and $(\varepsilon_i > s_{i,c_n}^{(m)})$ are equal, we find

$$1_{(\varepsilon_i/\sigma > c_n + n^{-1/4}A_1)} \le 1_{(y_i - x_i'\widehat{\beta}_{c_n}^{(m)} > \widehat{\sigma}_{c_n}^{(m)}c_n)} \le 1_{(\varepsilon_i/\sigma > c_n - n^{-1/4}A_1)}.$$

A similar argument shows

$$1_{(\varepsilon_i/\sigma < -c_n - n^{-1/4}A_1)} \le 1_{(y_i - x_i'\widehat{\beta}_{c_n}^{(m)} < -\widehat{\sigma}_{c_n}^{(m)}c_n)} \le 1_{(\varepsilon_i/\sigma < -c_n + n^{-1/4}A_1)}.$$

Thus, we get the lower and upper bound for indicators uniformly in iteration m so

$$1_{(|\varepsilon_i/\sigma| > c_n + n^{-1/4}A_1)} \le 1_{(|y_i - x_i'\widehat{\beta}_{c_n}^{(m)}| > \widehat{\sigma}_{c_n}^{(m)}c_n)} \le 1_{(|\varepsilon_i/\sigma| > c_n - n^{-1/4}A_1)}. \tag{43}$$

3. *Expectation of indicator bounds*. The aim is to prove

$$n\mathsf{E}1_{(|\varepsilon_i/\sigma| > c_n + n^{-1/4}A_1)} \to \lambda, \qquad n\mathsf{E}1_{(|\varepsilon_i/\sigma| > c_n - n^{-1/4}A_1)} \to \lambda. \tag{44}$$

Since $n\mathsf{E}1_{(|\varepsilon_i/\sigma| > c_n)} = \lambda$ by (15), it suffices to show

$$\mathscr{E}_n = n\mathsf{E}\{1_{(|\varepsilon_i/\sigma| > c_n - n^{-1/4}A_1)} - 1_{(|\varepsilon_i/\sigma| > c_n + n^{-1/4}A_1)}\} \to 0.$$

Note $|\varepsilon_i/\sigma| \sim \mathsf{g}$, G and $\mathsf{g} = 2\mathsf{f}$, $\mathsf{G} = 2\mathsf{F} - 1$. By this and (15), $2\{1 - \mathsf{F}(c_n)\} = \lambda/n$. Write $\mathscr{E}_n$ as integral, apply the mean value theorem and use the above identity to get

$$\mathscr{E}_n = n\int_{c_n - n^{-1/4}A_1}^{c_n + n^{-1/4}A_1} 2\mathsf{f}(u)du = 4nn^{-1/4}A_1\mathsf{f}(\tilde{c}) = \frac{4\lambda n^{-1/4}A_1\mathsf{f}(\tilde{c})}{2\{1 - \mathsf{F}(c_n)\}},$$

where $|\tilde{c} - c_n| \leq n^{-1/4} A_1$. Then, we find

$$\mathscr{E}_n = 2\lambda A_1 \frac{\mathsf{f}(\tilde{c})}{\mathsf{f}(c_n - n^{-1/4}A_1)} \frac{\mathsf{f}(c_n - n^{-1/4}A_1)}{\mathsf{f}(c_n)} \frac{\mathsf{f}(c_n)}{c_n\{1 - \mathsf{F}(c_n)\}} n^{-1/4} c_n.$$

Since $c_n - n^{-1/4}A_1 \leq \tilde{c}$ and f has the decreasing tail by Assumption 1(*ia*), the first ratio is bounded by 1. Since $c_n = \mathrm{o}(n^{1/4})$, Assumption 1(*ib*, *ic*) shows the second and third ratio are bounded. Then use $n^{-1/4}c_n = \mathrm{o}(1)$ to get $\mathscr{E}_n = \mathrm{o}(1)$.

4. *Poisson approximation*. On the set $\mathscr{B}_n$, apply (43) to obtain

$$\sum_{i=1}^{n} 1_{(|\varepsilon_i/\sigma| > c_n + n^{-1/4}A_1)} \leq \sum_{i=1}^{n} 1_{(|y_i - x_i'\widehat{\beta}_{c_n}^{(m)}| > \widehat{\sigma}_{c_n}^{(m)} c_n)} \leq \sum_{i=1}^{n} 1_{(|\varepsilon_i/\sigma| > c_n - n^{-1/4}A_1)}.$$

Using (44), the Poisson limit theorem shows that the lower and upper bound have the Poisson limit with mean λ. By (16), $n\widehat{\gamma}_{c_n}^{(m)} \stackrel{\mathsf{D}}{\to} \text{Poisson}(\lambda)$ for all $0 \leq m < \infty$. □

References

1. Bercu, B., Touati, A.: Exponential inequalities for self-normalized martingales with applications. Ann. Appl. Probab. **18**, 1848–1869 (2008)
2. Bickel, P.J.: One-step Huber estimates in the linear model. J. Am. Stat. Assoc. **70**, 428–434 (1975)
3. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)
4. Cavaliere, G., Georgiev, I.: Exploiting infinite variance through dummy variables in nonstationary autoregression. Econom. Theory **29**, 1162–1195 (2013)
5. Dollinger, M.B., Staudte, R.G.: Influence functions of iteratively re-weighted least squares estimators. J. Am. Stat. Assoc. **86**, 709–716 (1991)
6. Doornik, J.A.: Autometrics. In Castle, J.L., Shephard, N.: (eds.) The Methodology and Practice of Econometrics: A Festschrift in Honor of David F. Hendry, pp. 88–121. Oxford University Press, Oxford (2009)
7. Freedman, D.A.: On tail probabilities for martingales. Ann. Probab. **3**, 100–118 (1975)
8. Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A.: Robust Statistics. Wiley, New York (1986)
9. Hendry, D.F.: An econometric analysis of US food expenditure, 1931–1989. In: Magnus, J.R., Morgan, M.S. (eds.) Methodology and Tacit Knowledge: Two Experiments in Econometrics, pp. 341–361. Wiley, New York (1999)
10. Hendry, D.F., Doornik, J.A.: Empirical Model Discovery and Theory Evaluation. MIT Press, Cambridge MA (2014)
11. Hendry, D.F., Johansen, S., Santos, C.: Automatic selection of indicators in a fully saturated regression. Comput. Stat. **23**, 317–335 (2008). Erratum 337–339
12. Hendry, D.F., Santos, C.: An automatic test of super exogeneity. In Bollerslev, T., Russell, J.R., Watson, M.W. (eds.) Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle, pp. 164–193. Oxford University Press, Oxford (2010)
13. Hoover, K.D., Perez, S.J.: Data mining reconsidered: encompassing and the general-to-specific approach to specification search. Econom. J. **2**, 167–191 (1999)
14. Huber, P.J.: Robust estimation of a location parameter. Ann. Math. Stat. **35**, 73–101 (1964)

15. Johansen, S., Nielsen, B.: An analysis of the indicator saturation estimator as a robust regression estimator. In Castle, J.L., Shephard, N. (eds.) The Methodology and Practice of Econometrics: A Festschrift in Honor of David F. Hendry, pp. 1–36. Oxford University Press, Oxford (2009)
16. Johansen, S., Nielsen, B.: Asymptotic theory for iterated one-step Huber-skip estimators. Econometrics **1**, 53–70 (2013)
17. Johansen, S., Nielsen, B.: Analysis of the Forward Search using some new results for martingales and empirical processes. Bernoulli **22**, 1131–1183 (2016)
18. Johansen, S., Nielsen, B.: Asymptotic theory of outlier detection algorithms for linear time series regression models (with discussion). Scand. J. Stat. **43**, 321–381 (2016)
19. Jurečková, J., Sen, P.K., Picek, J.: Methodology in Robust and Nonparametric Statistics. CRC Press, Boca Raton (2013)
20. Koul, H.L.: Weighted Empirical Processes in Dynamic Nonlinear Models, 2nd edn. Springer, New York (2002)
21. Koul, H.L., Ling, S.: Fitting an error distribution in some heteroscedastic time series models. Ann. Stat. **34**, 994–1012 (2006)
22. Koul, H.L., Ossiander, M.: Weak convergence of randomly weighted dependent residual empiricals with applications to autoregression. Ann. Stat. **22**, 540–562 (1994)
23. Ruppert, D., Carroll, R.J.: Trimmed least squares estimation in the linear model. J. Am. Stat. Assoc. **75**, 828–838 (1980)
24. Varga, R.S.: Matrix Iterative Analysis. Springer, Berlin (2000)
25. Welsh, A.H., Ronchetti, E.: A journey in single steps: robust one-step M-estimation in linear regression. J. Stat. Plan. Inference **103**, 287–310 (2002)

Regression Quantile and Averaged Regression Quantile Processes

Jana Jurečková

Abstract We consider the averaged version $\widetilde{B}_n(\alpha)$ of the two-step regression α-quantile, introduced in [6] and studied in [7]. We show that it is asymptotically equivalent to the averaged version $\bar{B}_n(\alpha)$ of ordinary regression quantile and also study the finite-sample relation of $\widetilde{B}_n(\alpha)$ to $\bar{B}_n(\alpha)$. An interest of its own has the fact that the vector of slope components of the regression α-quantile coincides with a particular R-estimator of the slope components of regression parameter. Under a finite n, the stochastic processes $\widetilde{\mathscr{B}}_n = \{\widetilde{B}_n(\alpha) : 0 < \alpha < 1\}$ and $\mathscr{B}_n = \{\bar{B}_n(\alpha) : 0 < \alpha < 1\}$ differ only by a drift.

Keywords Averaged regression quantile · Regression quantile process · Two-step regression quantile process

1 Introduction

Consider the linear regression model

$$\mathbf{Y}_n = \mathbf{X}_n \beta + \mathbf{U}_n \tag{1}$$

with observations $\mathbf{Y}_n = (Y_1, \ldots, Y_n)^\top$, i.i.d. errors $\mathbf{U}_n = (U_1, \ldots, U_n)^\top$ with an unknown distribution function F, and unknown parameter $\beta = (\beta_0, \beta_1, \ldots, \beta_p)^\top$. The $n \times (p+1)$ matrix $\mathbf{X} = \mathbf{X}_n$ is known and $x_{i0} = 1$ for $i = 1, \ldots, n$ (i.e., β_0 is an intercept). The α-regression quantile $\widehat{\beta}_n(\alpha)$ of model (1) was introduced by Koenker and Bassett [10] as a solution of the minimization

$$\alpha \sum_{i=1}^{n} (Y_i - \mathbf{x}_i^\top \mathbf{b})^+ + (1-\alpha) \sum_{i=1}^{n} (Y_i - \mathbf{x}_i^\top \mathbf{b})^- := \min \tag{2}$$

J. Jurečková (✉)
Faculty of Mathematics and Physics, Department of Probability and Statistics, Charles University, Sokolovská 83, 186 75 Prague 8, Czech Republic
e-mail: jurecko@karlin.mff.cuni.cz

J. Antoch et al. (eds.), *Analytical Methods in Statistics*, Springer Proceedings in Mathematics & Statistics 193, DOI 10.1007/978-3-319-51313-3_3

with respect to $\mathbf{b} = (b_0, \ldots, b_p)^\top \in \mathbb{R}^{p+1}$, where $\mathbf{x}_i^\top$ is the i-th row of $\mathbf{X_n}$, $i = 1, \ldots, n$. The population counterpart of $\widehat{\beta}_n(\alpha)$ is the vector $\beta(\alpha) = (\beta_0 + F^{-1}(\alpha), \beta_1, \ldots, \beta_p)^\top$. For the brevity, we shall occasionally use the notation

$$\mathbf{x}_i^* = (x_{i1}, \ldots, x_{ip})^\top, \quad i = 1, \ldots, n \quad \text{and} \quad \mathbf{X}_n^* = \left[\mathbf{x}_1^*, \ldots, \mathbf{x}_n^*\right]^\top$$
$$\beta^* = (\beta_1, \ldots, \beta_p)^\top, \quad \widehat{\beta}_n^*(\alpha) = (\hat{\beta}_1(\alpha), \ldots, \hat{\beta}_p(\alpha))^\top.$$

The asymptotic behavior of α-regression quantile was studied in [2, 10, 14], among many others. The two-step regression α-quantile, asymptotically and numerically close to $\widehat{\beta}_n(\alpha)$, was proposed and studied in [6].

We impose the following regularity conditions on matrix $\mathbf{X}_n$ and on distribution function $F(x)$ of the errors U_i:

A1 $\lim_{n\to\infty} \mathbf{Q}_n = \mathbf{Q}$, where $\mathbf{Q}_n = n^{-1}\mathbf{X}_n^\top \mathbf{X}_n$ and $\mathbf{Q}$ is a positive definite matrix. The first column of $\mathbf{X}_n$ consists of ones.
A2 $n^{-1} \sum_{i=1}^n x_{ij}^4 = \mathscr{O}(1)$, as $n \to \infty$, for $j = 1, \ldots, p$.
B1 F has a continuous Lebesgue density f, which is positive and finite on $\{t : 0 < F(t) < 1\}$.

For two processes $\mathbf{A}_n$, $\mathbf{B}_n$ with realizations in $D^q(0, 1)$, we shall write

$$\mathbf{A}_n = \mathbf{B}_n + o_p^*(1) \;\; [\text{ or } O_p^*(1), o^*(1), O^*(1), \;\; \text{respectively}], \quad \text{if}$$

$$\begin{aligned}\|\mathbf{A}_n - \mathbf{B}_n\|_\varepsilon &= \sup_{\varepsilon \le \alpha \le 1-\varepsilon} \|\mathbf{A}_n(\alpha) - \mathbf{B}_n(\alpha)\| \\ &= o_p(1) \; [\text{ or } \; \mathrm{O_p}(1), \mathrm{o}(1), \mathrm{O}(1), \; \text{respectively}]\end{aligned}$$

for all $\varepsilon \in (0, 1/2)$. If $\mathbf{B}_n \xrightarrow{\mathscr{D}} \mathbf{B}$ and $\mathbf{A}_n = \mathbf{B}_n + o_p^*(1)$, then $\mathbf{A}_n \xrightarrow{\mathscr{D}} \mathbf{B}$.

Let $\mathbf{W}_n^{\mathbf{x}}$ denotes the weighted empirical process

$$\mathbf{W}_n^{\mathbf{x}} = n^{-1/2} \sum_{i=1}^n \mathbf{x}_{ni}(a_i^*(\alpha) - (1-\alpha)), \quad \text{where } \; a_i^*(\alpha) = I[U_i > F^{-1}(\alpha)]$$

$i = 1, \ldots, n$, and consider the regression quantile process

$$\begin{aligned}\mathbf{Z}_n &= \left\{\mathbf{Z}_n(\alpha) = n^{1/2}(\widehat{\boldsymbol{\beta}}_n(\alpha) - \beta(\alpha)), \; 0 < \alpha < 1\right\}, \\ \beta(\alpha) &= \beta + F^{-1}(\alpha)\mathbf{e}_1, \; \mathbf{e}_1 = (1, 0, \ldots, 0)^\top \in \mathbb{R}^{p+1}.\end{aligned} \quad (3)$$

For the sake of completeness, let $Z_n^{(0)}$ denote the ordinary quantile process:

$$Z_n^{(0)} = \left\{ Z_n^{(0)}(\alpha) = n^{1/2}(F_n^{-1}(\alpha) - F^{-1}(\alpha)),\ 0 < \alpha < 1 \right\} \tag{4}$$

where F_n^{-1} is the empirical quantile function.

A uniform asymptotic representation of $\mathbf{Z}_n$ up to the remainder term of order $0(n^{-1/4}\log n)$ was derived in [13]. The Bahadur-type representation of $\mathbf{Z}_n$ under weaker conditions, and its convergence to a Gaussian process was proven in [2]. More precisely,

Theorem 1 *Under the conditions* **A1** *and* **A2***,*

$$\mathbf{Z}_n(\alpha) = \frac{1}{f(F^{-1}(\alpha))}\mathbf{Q}_n^{-1}\sum_{i=1}^{n}\mathbf{W}_n^{\mathbf{x}}(\alpha) + o_p^*(1) \tag{5}$$

as $n \to \infty$.

Moreover,

$$\mathbf{Z}_n \overset{\mathscr{D}}{\to} \left(f \circ F^{-1}\right)^{-1}\mathbf{Q}^{-1}\mathbf{W}_{(p)}^* \quad \textit{as} \ \ n \to \infty, \tag{6}$$

where $\mathbf{W}_{(p)}^*$ *is a vector of* p *independent Brownian bridges on* $(0, 1)$.

The weighted empirical processes of type (5) and their asymptotic properties were systematically studied in Koul's monograph [11] with a rich bibliography.

2 Averaged Regression Quantile Process

The regression quantiles were widely applied in the statistical and econometric inference; here we refer to Koenker's monograph [9] and to the references cited in, among others. Their extension to the autoregression processes was studied in [12].

The scalar statistic

$$\bar{B}_n(\alpha) = \bar{\mathbf{x}}_n^\top \widehat{\beta}_n(\alpha), \qquad \bar{\mathbf{x}}_n = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{ni} \tag{7}$$

is called the *averaged regression* α*-quantile*; it was first considered in [1]. It is in $\alpha \in (0, 1)$ a.s., and scale and regression equivariant in the sense that

$$\bar{B}_n(\alpha; \mathbf{Y} + \mathbf{X}\mathbf{b}) = \bar{B}_n(\alpha, \mathbf{Y}) + \bar{\mathbf{x}}^\top\mathbf{b} \quad \forall \mathbf{b} \in \mathbb{R}^{p+1}.$$

For every fixed n, $\bar{B}_n(\alpha)$ equals to a linear combination of $p + 1$ components of the vector of responses $\mathbf{Y}$, corresponding to the optimal base of the linear programming, leading to calculation of $\widehat{\beta}_n(\alpha)$ or of its dual.

As it was shown in [8], $\bar{B}_n(\alpha)$ is asymptotically equivalent to the $[n\alpha]$-quantile of the location model. More precisely, under the conditions of Theorem 1,

$$n^{1/2}\left[\bar{\mathbf{x}}_n^\top(\widehat{\beta}_n(\alpha)-\widetilde{\beta}(\alpha))-(U_{n:[n\alpha]}-F^{-1}(\alpha))\right]=\mathscr{O}_p(n^{-1/4}) \tag{8}$$

as $n\to\infty$, where $U_{n:1}\le\cdots\le U_{n:n}$ are the order statistics corresponding to $U_1,\dots,U_n$.

Let us now consider *the average regression quantile process*

$$\bar{\mathscr{B}}_n=\left\{n^{1/2}\bar{\mathbf{x}}_n^\top\Big(\widehat{\beta}_n(\alpha)-\beta(\alpha)\Big);\ 0<\alpha<1\right\}.$$

Its trajectories are step functions, nondecreasing in $\alpha\in(0,1)$, and for each n they have finite numbers of discontinuities. As was shown in [1],

$$\bar{B}_n(\alpha_1)\neq\bar{B}_n(\alpha_2)\ \text{ with probability 1, provided }\ 1>\alpha_2-\alpha_1>\frac{p+1}{n}>0.$$

Combining (8) with Theorem 1, we come to the conclusion that the process $\bar{\mathscr{B}}_n$ is asymptotically equivalent to the location quantile process, and that it converges to a Gaussian process in the Skorokhod topology. More precisely,

Theorem 2 *Under the conditions **A1**, **A2**, and **B1**, the process $\bar{\mathscr{B}}_n$ admits the asymptotic representation*

$$\bar{\mathscr{B}}_n=n^{-1/2}\frac{1}{f(F^{-1}(\alpha))}\sum_{i=1}^n\Big(I[U_i>F^{-1}(\alpha)]-(1-\alpha)\Big)+o_p^*(1). \tag{9}$$

Moreover,

$$\bar{\mathscr{B}}_n\xrightarrow{\mathscr{D}}(f\circ F^{-1})^{-1}W^*\ \text{ as }\ n\to\infty \tag{10}$$

where W^* is the Brownian bridge on (0,1).

The weak convergence of process $\bar{\mathscr{B}}_n$ extends to various functionals, what has useful applications, e.g., in models with nuisance regression. The representation (9) coincides with the representation of the ordinary quantile process (4), hence $\bar{\mathscr{B}}_n$ is asymptotically equivalent to the ordinary quantile process. Moreover, $\bar{B}_n(0)\le\bar{B}_n(\alpha)\le\bar{B}_n(1)$, where $\bar{B}_n(0)$ and $\bar{B}_n(1)$ are extreme averaged regression quantiles, studied in [5]. As such, the inversion of $\{\bar{B}_n(\alpha),\ 0<\alpha<1\}$

$$H_n(z)=\inf\{\alpha:\ \bar{B}_n(\alpha)\ge z,\ z\in\mathbb{R}\}$$

makes a sense. It is again a nondecreasing step function with a finite number of discontinuities, $H_n(-\infty)=0,\ H_n(\infty)=1$, hence it is a discrete distribution function on $\mathbb{R}^1$. Portnoy [13] demonstrated its tightness with respect to the Skorokhod topology and that of its smoothed version in $\mathscr{C}(0,1)$. We can conclude that $n^{1/2}(H_n(F^{-1}(\alpha))-\alpha)$ converges weakly to the Brownian bridge W^* on $(0,1)$.

In the following section, we shall parallely consider the two-step regression α-quantile $\widetilde{\beta}_n(\alpha)$, introduced in [6], and its averaged version $\widetilde{B}_n(\alpha)$; they are asymptotically equivalent to $\widehat{\beta}_n(\alpha)$ and to $\bar{B}_n(\alpha)$, respectively. $\widetilde{B}_n(\alpha)$ can be equivalently expressed as the $[n\alpha]$-th order statistic of residuals of Y_i's from specific R-estimators of the slope components; this enables to look more closely in the structure of the considered concepts. Moreover, we shall also describe the finite sample relation of $\widetilde{B}_n(\alpha)$ to $\widehat{B}_n(\alpha)$.

For every fixed n, $\bar{B}_n(\alpha)$ equals to a linear combination of $p+1$ components of the vector of responses $\mathbf{Y}$, corresponding to the optimal base of the linear programming, leading to calculation of $\widehat{\beta}_n(\alpha)$ or of its dual.

Let us now consider *the average regression quantile process*

$$\bar{\mathscr{B}}_n = \left\{ n^{1/2} \bar{\mathbf{x}}_n^\top \left(\widehat{\beta}_n(\alpha) - \widetilde{\beta}(\alpha) \right);\ 0 < \alpha < 1 \right\}.$$

Its trajectories are step functions, nondecreasing in $\alpha \in (0, 1)$, and for each n they have finite numbers of discontinuities. As was shown in [1],

$$\bar{B}_n(\alpha_1) \neq \bar{B}_n(\alpha_2) \ \text{ with probability 1, provided } \ 1 > \alpha_2 - \alpha_1 > \frac{p+1}{n} > 0.$$

Combining (8) with Theorem 1, we come to the conclusion that the process $\bar{\mathscr{B}}_n$ is asymptotically equivalent to the location quantile process, and that it converges to a Gaussian process in the Skorokhod topology. More precisely,

Theorem 3 *Under the conditions **A1**, **A2** and **B1**, the process $\bar{\mathscr{B}}_n$ admits the asymptotic representation*

$$\bar{\mathscr{B}}_n = n^{-1/2} \frac{1}{f(F^{-1}(\alpha))} \sum_{i=1}^{n} \Big(I[U_i > F^{-1}(\alpha)] - (1-\alpha) \Big) + o_p^*(1). \tag{11}$$

Moreover,

$$\bar{\mathscr{B}}_n \xrightarrow{\mathscr{D}} (f \circ F^{-1})^{-1} W^* \ \ as \ \ n \to \infty \tag{12}$$

where W^* is the Brownian bridge on (0,1).

The weak convergence of process $\bar{\mathscr{B}}_n$ extends to various functionals, what has useful applications, e.g. in models with nuisance regression. The representation (9) coincides with the representation of the ordinary quantile process (4), hence $\bar{\mathscr{B}}_n$ is asymptotically equivalent to the ordinary quantile process. Moreover, $\bar{B}_n(0) \leq \bar{B}_n(\alpha) \leq \bar{B}_n(1)$, where $\bar{B}_n(0)$ and $\bar{B}_n(1)$ are extreme averaged regression quantiles, studied in [5]. As such, the inversion of $\{\bar{B}_n(\alpha),\ 0 < \alpha < 1\}$

$$H_n(z) = \inf\{\alpha : \ \bar{B}_n(\alpha) \geq z,\ z \in \mathbb{R}\}$$

makes a sense. It is again a nondecreasing step function with a finite number of discontinuities, $H_n(-\infty) = 0,\ H_n(\infty) = 1$, hence it is a discrete distribution function on $\mathbb{R}^1$. Portnoy [13] demonstrated its tightness with respect to the Skorokhod topology and that of its smoothed version in $\mathscr{C}(0,1)$. We can conclude that $n^{1/2}(H_n(F^{-1}(\alpha)-\alpha)$ converges weakly to the Brownian bridge W^* on (0,1).

In the following section, we shall parallely consider the two-step regression α-quantile $\widetilde{\beta}_n(\alpha)$, introduced in [6], and its averaged version $\widetilde{B}_n(\alpha)$; they are asymptotically equivalent to $\widehat{\beta}_n(\alpha)$ and to $\bar{B}_n(\alpha)$, respectively. $\widetilde{B}_n(\alpha)$ can be equivalently expressed as the $[n\alpha]$-th order statistic of residuals of Y_i's from specific R-estimators of the slope components; this enables to look more closely in the structure of the considered concepts. Moreover, we shall also describe the final sample relation of $\widetilde{B}_n(\alpha)$ to $\widehat{B}_n(\alpha)$.

3 Averaged Two-Step Regression Quantile Process

The slope component of the two-step regression α-quantile $\widetilde{\beta}_n(\alpha)$ is a specific R-estimate $\widetilde{\beta}^*_{nR}(\alpha)$ of the slope parameter vector β^*; the intercept component is the $[n\alpha]$-quantile of residuals of Y_i's from $\widetilde{\beta}^*_{nR}(\alpha)$. The result is a consistent estimator of $(\beta_0+F^{-1}(\alpha),\beta_1,\ldots,\beta_p)^\top$, is asymptotically equivalent to the regression α-quantile $\widehat{\beta}_n(\alpha)$, and very close to the same even for finite samples.

Consider the model (1) under the following conditions on F and on $\mathbf{X}_n$:

(A1) F has a continuous density f that is positive on the support of F and has finite Fisher's information, i.e.

$$0<\int\left(\frac{f'(x)}{f(x)}\right)^2 dF(x)<\infty.$$

(A2) (*Generalized Noether condition.*) $\lim_{n\to\infty}\max_{1\le i\le n}\mathbf{x}^{*\top}_{ni}\left(\sum_{k=1}^n \mathbf{x}^*_{nk}\mathbf{x}^{*\top}_{nk}\right)^{-1}\mathbf{x}^*_{ni}=0.$

(A3) $\lim_{n\to\infty} n^{-1}\sum_{i=1}^n \mathbf{x}_{ni}\mathbf{x}_{ni}^\top=\mathbf{Q}$, where $\mathbf{x}_{ni}=(1,x^*_{n,i1},\ldots,\mathbf{x}^*_{n,ip})^\top$, $i=1,\ldots,n$, and $\mathbf{Q}$ is a positively definite $(p+1)\times(p+1)$ matrix.

Let $\widetilde{\beta}^*_{nR}(\alpha)$ be the R-estimator of the slope components $\beta^*=(\beta_1,\ldots,\beta_p)^\top$ based on the score-generating function

$$\varphi_\alpha(u)=\alpha-I[u<\alpha],\ 0<u<1. \tag{13}$$

It is a minimizer of Jaeckel's [4] measure of rank dispersion

$$\begin{aligned}\widetilde{\beta}_{nR}(\alpha)&=\operatorname{argmin}_{\mathbf{b}\in\mathbb{R}^p}\mathscr{D}_n(\mathbf{b}),\\ \mathscr{D}_n(\mathbf{b})&=\sum_{i=1}^n(Y_i-\mathbf{x}_i^{*\top}\mathbf{b})\varphi_\alpha\left(\frac{R_{ni}(Y_i-\mathbf{x}_i^{*\top}\mathbf{b})}{n+1}\right),\end{aligned} \tag{14}$$

where $R_{ni}(Y_i-\mathbf{x}_i^{*\top}\mathbf{b})$ is the rank of the i-th residual, $i=1,\ldots,n$.

Having estimated β^* by R-estimate $\widetilde{\beta}^*_{nR}(\alpha)$, we define the intercept component $\tilde{\beta}_{n0}(\alpha)$ of the two-step regression quantile as the $[n\alpha]$-order statistic of the residuals $Y_i - \mathbf{x}_i^{*\top}\widetilde{\beta}^*_{nR}(\alpha),\ i = 1, \ldots, n$. The two-step α-regression quantile is then the vector

$$\widetilde{\beta}_n(\alpha) = \begin{pmatrix} \tilde{\beta}_{n0}(\alpha) \\ \widetilde{\beta}^*_{nR}(\alpha) \end{pmatrix} \in \mathbb{R}^{p+1}. \tag{15}$$

As shown in [6], $\widetilde{\beta}_n(\alpha)$ is asymptotically equivalent to the Koenker–Bassett α-regression quantile, uniformly over $(\varepsilon, 1-\varepsilon) \subset (0,1),\ \forall \varepsilon \in (0, \frac{1}{2})$, hence

$$n^{\frac{1}{2}} \|\widehat{\beta}_n(\alpha) - \widetilde{\beta}_n(\alpha)\| = o^*_p(1) \text{ as } n \to \infty. \tag{16}$$

We shall show that, under finite n, the slope components of $\widetilde{\beta}_n(\alpha)$ and of $\widehat{\beta}_n(\alpha)$ exactly coincide for every fixed $\alpha \in (0,1)$. It is expressed in the following lemma:

Lemma 1 *Let $\widehat{\beta}_n(\alpha)$ and $\widetilde{\beta}_n(\alpha)$ be the α-regression quantile and two-step α-regression quantile in model (1), $\alpha \in (0,1)$. Then, their slope components coincide, i.e., $\widehat{\beta}^*_n(\alpha) \equiv \widetilde{\beta}^*_{nR}(\alpha)$ for every fixed $\alpha \in (0,1)$.*

Proof The minimum in (14), being attained for $\widetilde{\beta}^*_{nR}(\alpha)$, can be rewritten as

$$\sum_{i=1}^{n} (Y_i + \bar{\mathbf{x}}_n^\top \widetilde{\beta}^*_{nR}(\alpha) - \mathbf{x}_i^{*\top}\widetilde{\beta}^*_{nR}(\alpha))[a_i(\alpha, \widetilde{\beta}^*_{nR}(\alpha)) - (1-\alpha)] \tag{17}$$

with respect to $\mathbf{b} \in \mathbb{R}^p$, where for $\mathbf{b} \in \mathbb{R}^p$

$$a_i(\alpha, \mathbf{b}) = \begin{cases} 0 & \ldots \quad R_{ni}(Y_i - \mathbf{x}_i^{*\top}\mathbf{b}) < n\alpha \\ R_i - n\alpha & \ldots \quad n\alpha \le R_{ni}(Y_i - \mathbf{x}_i^{*\top}\mathbf{b}) < n\alpha + 1 \\ 1 & \ldots \quad n\alpha + 1 \le R_{ni}(Y_i - \mathbf{x}_i^{*\top}\mathbf{b}). \end{cases} \tag{18}$$

The $a_i(\alpha, \mathbf{b})$ are known as *Hájek's* [3] *rank scores*. Remind that they are invariant to the translation and satisfy

$$\sum_{i=1}^{n} a_i(\alpha, \mathbf{b}) = n(1-\alpha).$$

Moreover, due to the linear programing duality to minimization (2) for location model and $Y_i \mapsto Y_i - \mathbf{x}_i^{*\top}\mathbf{b},\ i = 1, \ldots, n,$ vector $\mathbf{a}_n(\alpha, \mathbf{b}) = (a_1(\alpha, \mathbf{b}), \ldots, a_n(\alpha, \mathbf{b}))^\top$ solves the maximization

$$\sum_{i=1}^{n} (Y_i - \mathbf{x}_i^{*\top}\mathbf{b})[a_i(\alpha, \mathbf{b}) - (1-\alpha)] = \max$$

under the constraint

$$\sum_{i=1}^{n} a_i(\alpha, \mathbf{b}) = n(1-\alpha),\ 0 \le a_i(\alpha, \mathbf{b}) \le 1,\ i = 1, \ldots, n.$$

Hence, combining the minimization with respect to $\mathbf{b} \in \mathbb{R}^p$ and maximization with respect to $0 \le a_i(\alpha, \mathbf{b}) \le 1,\ i = 1, \ldots, n,\ \sum_{i=1}^{n} a_i(\alpha, \mathbf{b}) = n(1-\alpha)$, we obtain for a fixed α

$$\begin{aligned}
&\alpha \sum_{i=1}^{n} (Y_i - \mathbf{x}_i^\top \widehat{\beta}_n(\alpha))^+ + (1-\alpha) \sum_{i=1}^{n} (Y_i - \mathbf{x}_i^\top \widehat{\beta}_n(\alpha))^- \\
&= \sum_{i=1}^{n} Y_i [\hat{a}_i(\alpha) - (1-\alpha)] \qquad (19) \\
&= \sum_{i=1}^{n} (Y_i + \bar{\mathbf{x}}_n^{*\top} \widetilde{\beta}_{nR}^*(\alpha) - \mathbf{x}_i^{*\top} \widetilde{\beta}_{nR}^*(\alpha))[\hat{a}_i(\alpha) - (1-\alpha)] \\
&\le \sum_{i=1}^{n} (Y_i + \bar{\mathbf{x}}_n^{*\top} \widetilde{\beta}_{nR}^*(\alpha) - \mathbf{x}_i^{*\top} \widetilde{\beta}_{nR}^*(\alpha))[a_i(\alpha, \widetilde{\beta}_{nR}^*(\alpha)) - (1-\alpha)] \\
&\le \sum_{i=1}^{n} (Y_i - \mathbf{x}_i^\top \widehat{\beta}_n(\alpha))[a_i(\alpha, \widetilde{\beta}_{nR}^*(\alpha)) - (1-\alpha)] \\
&\le \sum_{i=1}^{n} (Y_i - \mathbf{x}_i^\top \widehat{\beta}_n(\alpha))[a_i(\alpha, \widehat{\beta}_n^*(\alpha)) - (1-\alpha)]
\end{aligned}$$

where $\hat{a}_i(\alpha),\ i = 1, \ldots, n$ are regression rank scores, dual to $\widehat{\beta}_n(\alpha)$, defined as a solution of the linear program

$$\begin{aligned}
&\sum_{i=1}^{n} Y_i\ \hat{a}_i(\alpha) := \max \\
\text{under the constraint}\quad &\sum_{i=1}^{n} \hat{a}_i(\alpha) = n(1-\alpha), \\
&\sum_{i=1}^{n} x_{ij} \hat{a}_i(\alpha) = (1-\alpha) \sum_{i=1}^{n} x_{ij},\ j = 1, \ldots, p, \\
&0 \le \hat{a}_i(\alpha) \le 1, \quad i = 1, \ldots, n, \quad 0 \le \alpha \le 1.
\end{aligned}$$

Let $\hat{r}_i(\alpha) = Y_i - \mathbf{x}_i^\top \widehat{\beta}_n(\alpha),\ i = 1, \ldots, n$ and $\hat{r}_{n:1}(\alpha) \le \hat{r}_{n:2}(\alpha) \le \ldots \le \hat{r}_{n:n}(\alpha)$. Then the last term in (19) can be rewritten as

$$(\alpha - 1) \sum_{i=1}^{\lceil n\alpha \rceil - 1} \hat{r}_{n:\lceil n\alpha \rceil}(\alpha) + \alpha \sum_{i=\lceil n\alpha \rceil}^{n} \hat{r}_{n:i}(\alpha). \tag{20}$$

Consider the difference of the first term in (19) and of (20), which is nonpositive. If $\hat{r}_{n:\lceil n\alpha \rceil}(\alpha) > 0$, then this difference equals to

$$\sum_{i=1}^{n} \hat{r}_i(\alpha) I[0 < \hat{r}_i(\alpha) < \hat{r}_{n:\lceil n\alpha \rceil}(\alpha)] \geq 0.$$

Similarly, if $\hat{r}_{n:\lceil n\alpha \rceil}(\alpha) < 0$, then the difference equals

$$-\sum_{i=1}^{n} \hat{r}_i(\alpha) I[\hat{r}_{n:\lceil n\alpha \rceil}(\alpha) < \hat{r}_i(\alpha) < 0] \geq 0$$

and the difference equals 0 otherwise. Hence, the first and the last terms, and also all terms in (19) coincide. This further implies that $\widehat{\beta}_n^*(\alpha)$ minimizes (17) and equals to $\widetilde{\beta}_{nR}^*(\alpha)$. □

Define the *averaged two-step regression α-quantile* $\widetilde{B}_n(\alpha)$, analogous to $\bar{B}_n(\alpha)$ in (7), as

$$\widetilde{B}_n(\alpha) = \bar{\mathbf{x}}_n^\top \widetilde{\beta}_n(\alpha), \qquad \bar{\mathbf{x}}_n = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_{ni} \tag{21}$$

$\widetilde{B}_n(\alpha)$ is scale equivariant and regression equivariant generally not monotone in α. By (15), for every fixed $\alpha \in (0, 1)$ is $\widetilde{B}_n(\alpha)$ equal to the $[n\alpha]$-order statistic of the residuals

$$\widetilde{B}_n(\alpha) = \left(Y_{ni} - (\mathbf{x}_{ni} - \bar{\mathbf{x}}_n)^\top \widetilde{\beta}_{nR}^*(\alpha)\right)_{n:[n\alpha]}. \tag{22}$$

The following theorem describes the finite-sample relation of $\widetilde{B}_n(\alpha)$ to $\bar{B}_n(\alpha)$:

Theorem 4 *Let $\bar{B}_n(\alpha)$ and $\widetilde{B}_n(\alpha)$ be the averaged regression α-quantile and the averaged two-step regression α-quantile, respectively. Then, for every fixed $\alpha \in (0, 1)$*

$$\widetilde{B}_n(\alpha) = \bar{B}_n(\alpha) + \left(Y_{ni} - \mathbf{x}_{ni}^\top \widehat{\beta}_n(\alpha)\right)_{n:[n\alpha]}. \tag{23}$$

Proof By (22) and by Lemma 1

$$\begin{aligned}
\widetilde{B}_n(\alpha) &= \left(Y_{ni} - (\mathbf{x}_{ni} - \bar{\mathbf{x}}_n)^\top \widetilde{\beta}_{nR}^*(\alpha)\right)_{n:[n\alpha]} \\
&= \left(Y_{ni} - \hat{\beta}_0(\alpha) - \mathbf{x}_{ni}^{*\top} \widetilde{\beta}_n^*(\alpha)\right)_{n:[n\alpha]} + \hat{\beta}_0(\alpha) + \bar{\mathbf{x}}_n^{*\top} \widehat{\beta}_n^*(\alpha) \\
&= \left(Y_{ni} - \mathbf{x}_{ni}^\top \widehat{\beta}_n(\alpha)\right)_{n:[n\alpha]} + \bar{\mathbf{x}}_n^\top \widehat{\beta}_n(\alpha) = \bar{B}_n(\alpha) + \left(Y_{ni} - \mathbf{x}_{ni}^\top \widehat{\beta}_n(\alpha)\right)_{n:[n\alpha]}.
\end{aligned}$$ □

Corollary 1 *Let $\mathscr{B}_n^1$ and $\mathscr{B}_n^2$ be the processes*

$$\mathscr{B}_n^1 = \{\bar{B}_n(\alpha) : 0 < \alpha < 1\}$$
$$\mathscr{B}_n^2 = \{\widetilde{B}_n(\alpha) : 0 < \alpha < 1\}.$$

Then

$$\mathscr{B}_n^2(\alpha) - \mathscr{B}_n^1(\alpha) - \big(Y_{ni} - \mathbf{x}_{ni}^\top \widehat{\beta}_n(\alpha)\big)_{n:[n\alpha]} = o_p^*(1) \quad \text{as } n \to \infty. \tag{24}$$

Corollary 2 *Let $\hat{\beta}_{n0}(\alpha)$ and $\tilde{\beta}_{n0}(\alpha)$ be the intercept components of the α-regression quantile and of the α-two-step regression quantile, respectively. Then for every fixed $\alpha \in (0, 1)$*

$$\tilde{\beta}_{n0}(\alpha) - \hat{\beta}_{n0}(\alpha) = \Big(Y_i - \mathbf{x}_{ni}^\top \widehat{\beta}_n(\alpha)\Big)_{n:[n\alpha]}. \tag{25}$$

Remark 1 The asymptotic relation of the α-regression quantile and of the α-two step regression quantile is numerically illustrated in [6].

Acknowledgements The author's research was supported by the Czech Republic Grant 15-00243S.

References

1. Bassett, G.W., Koenker, R.W.: An empirical quantile function for linear models with iid errors. J. Am. Stat. Assoc. **77**, 405–415 (1982)
2. Gutenbrunner, C., Jurečková, J.: Regression rank scores and regression quantiles. Ann. Stat. **20**, 305–330 (1992)
3. Hájek, J.: Extensions of the Kolmogorov–Smirnov tests to regression alternatives. In: LeCam, L. (ed.) Bernoulli–Bayes–Laplace Seminar, pp. 45–60 (1965)
4. Jaeckel, L.A.: Estimating regression coefficients by minimizing the dispersion of the residuals. Ann. Math. Stat. **43**, 1449–1459 (1972)
5. Jurečková, J.: Finite sample behavior of averaged extreme regression quantile. EXTREMES **19**, 41–49 (2016)
6. Jurečková, J., Picek, J.: Two-step regression quantiles. Sankhya 67. Part 2, pp. 227–252 (2005)
7. Jurečková, J., Picek, J.: Regression quantiles and their two-step modifications. Stat. Probab. Lett. **82**, 1111–1115 (2012)
8. Jurečková, J., Picek, J.: Averaged regression quantiles. In: Lahiri, S.N., et al. (eds.) *Contemporary Developments in Statistical Theory*, Springer Proceedings in Mathematics & Statistics, vol. 68, Chapter 12, pp. 203–216 (2014)
9. Koenker, R.: Quantile Regression. Cambridge University Press, Cambridge (2005)
10. Koenker, R., Bassett, G.: Regression quantiles. Econometrica **46**, 33–50 (1978)
11. Koul, H.L.: Weighted Empirical Processes in Dynamic Nonlinear Models. Lecture Notes in Statistics, vol. 166. Springer, Berlin (2002)
12. Koul, H.L., Saleh, AKMdE: Autoregression quantiles and related rank-scores processes. Ann. Stat. **23**, 670–689 (1995)
13. Portnoy, S.: Tightness of the sequence of empiric c.d.f. processes defined from regression fractiles. In: Franke, J., Haerdle, W., Martin, D. (eds.) Robust and Nonlinear Time Series Analysis. Lecture Notes in Statistics, vol. 26, pp. 231–246. Springer, New York (1984)
14. Ruppert, D., Carroll, R.J.: Trimmed least squares estimation in the linear model. J. Am. Stat. Assoc. **75**, 828–838 (1980)

Stability and Heavy-Tailness

Lev B. Klebanov

Abstract We discuss some simple statistical models leading to some families of probability distributions. These models are of specific interest because the desirable statistical property leads to functional equations having a large set of solutions. It appears that a small subset only of the set of all the solutions has probabilistic sense.

Keywords Characterization problems · ν-stable distributions · Heavy-tailed distributions

1 Introduction: A Little Bit of Naive Philosophy

When we were students, I heard the sentence that Probability is a part of measure theory, studying special case of positive normalized measure and a notion of independence. But what are typical results, showing the difference of Probability from Analysis? To me, typical results of such kind are Theorems by Cramér, Raikov, and Linnik on decompositions of Normal, Poisson, and the composition of Normal and Poisson distributions. These theorems provide us with very nice probability results whose proofs essentially use the analytic functions theory. They give also examples, which are not typical for classical Analysis. Namely, without positiveness property of the measures such results are impossible. Do we have analytical results typical for Mathematical Statistics?

2 Polya Theorem

I think, such results arise in Characterization of Probability distributions. The aim of Characterization is to describe all distributions of random variables possessing a

L.B. Klebanov (✉)
Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Charles University, Sokolovská 49/83, 18675 Prague, Czech Republic
e-mail: Lev.Klebanov@mff.cuni.cz

J. Antoch et al. (eds.), *Analytical Methods in Statistics*, Springer Proceedings in Mathematics & Statistics 193, DOI 10.1007/978-3-319-51313-3_4

desirable property, which may be taking as a base of probabilistic and/or statistical model. Let us start with an example leading to Polya theorem. Suppose that we have a gas whose molecules are chaotically moving, and the space is isotropic and homogeneous. Denote by X_1 and X_2 projections of the velocity of a molecule on the axis in $x - y$ plain. In view of space property we have the following properties: (a) X_1 and X_2 are independent random variables; (b) $X_1 \stackrel{d}{=} X_2$. After rotation of the coordinate system counterclockwise on the angle $\pi/4$ we obtain, that a projection on new coordinate axes has to be identically distributed with the old one. That is, $X_1 \stackrel{d}{=} (X_1 + X_2)/\sqrt{2}$. *Polya Theorem says that in this situation X_1 has normal (or degenerate) distribution with zero mean.*

Let us note that there are no moment conditions in Polya Theorem. This Theorem states that such general properties of the space like isotropy and homogeneity imply normality of corresponding velocity distribution. This model contains one parameter (the variance of Gaussian distribution), and a statistician has to make some inferences on this parameter. Polya theorem is a sophisticated result in the sense that it uses essentially the probabilist character of corresponding measures. Actually, to prove this theorem it is necessary to solve the following functional equation

$$f(t) = f^2(t/\sqrt{2}),$$

where $f(t)$ is characteristic function of X_1. The general solution of this equation has the form

$$f(t) = \exp\{-at^2 h(t)\},$$

where $h(\log t)$ is arbitrary continuous $(\log 2)/2$-periodic function. The probabilist character $f(t)$ allows to prove that $h(t) = const$ and $a \geq 0$.

3 ν-normality and ν-stability

Suppose that $U = \{X, X_1, \ldots, X_n, \ldots\}$ is a sequence of i.i.d. random variables. Let Δ be a subset of $(0, 1)$. Let $\{\nu_p,\ p \in \Delta\}$ be a family of positive integer-valued random variables independent with U. Let $p = 1/\mathbb{E}\nu_p < 1$ for all $p \in \Delta$. We say, X has ν-strictly stable distribution if there exist positive α such that

$$X \stackrel{d}{=} p^{1/\alpha} \sum_{k=1}^{\nu_p} X_k$$

for all $p \in \Delta$. The definition in this form and description of all ν-strictly stable distribution belongs to L.B. Klebanov and S.T. Rachev [7]. The particular case of geometrically distributed ν_p was considered by L.B. Klebanov, G.M. Manija and I.A. Melamed [5]. The same authors gave a generalization for the case of other ν_p under additional conditions in 1987 (see, [6]).

For the case of $\alpha = 2$, we are talking about ν-Gaussian (or ν-normal) distribution, and for the case of positive random variables and $\alpha = 1$ about ν-degenerate distribution (but they are nondegenerate as probability distributions).

Let us note that stable distributions in classical case (that is for the case of non-random number of summands) may be defined in two ways. The first one uses the property of identical distribution of a random variable with normalized sum of its i.i.d. copies (algebraic definition). The second way consists in defining stable distributions as limit laws for sums of i.i.d. random variables (analytic or limit definition). For the case of random number of summands, these definitions appears to be different. Given definition has algebraic character. Limit definition (which is now different with algebraic one) belongs to B.V. Gnedenko [2, 3].

The condition for existence of algebraic ν-strictly stable distribution has the following form. *Let $\mathfrak{G}$ be a semigroup, generated by all probability generating functions of $\{\nu_p,\ p \in \Delta\}$ with superposition as semigroup operation. ν-strictly stable distribution exists if and only if $\mathfrak{G}$ is commutative*. Let us note that notion and properties of ν-stable distributions are identical for the case of the family $\{\nu_p,\ p \in \Delta\}$ and for that of the semigroup, generated by this family. However, the sets Δ for these cases may be different. For example, the set Δ connected to $\mathfrak{G}$ always contains zero as a limit point. Everywhere below we suppose, that the set Δ is connected with the semigroup $\mathfrak{G}$.

An analogue of Polya theorem may be formulated in the following way. Suppose that the semigroup $\mathfrak{G}$ is commutative. *Suppose that U and $\{\nu_p,\ p \in \Delta\}$ are as above. The equality in distribution*

$$X \stackrel{d}{=} \sqrt{p} \sum_{k=0}^{\nu_p} X_k$$

is true for a fixed $p \in \Delta$ if and only if X has *ν-Gaussian distribution*. This theorem is also sophisticated in the sense that corresponding functional equation has infinitely many nonprobabilistic solutions.

In the case of

$$X \stackrel{d}{=} p \sum_{k=0}^{\nu_p} X_k$$

and positive random variables, we have similar characterization of ν-degenerate distribution.

In a very natural way one can define ν-infinite divisible distributions.

Let $\varphi(s)$ is Laplace transform of ν-degenerate random variable (see, for Example, [8]). *Characteristic function $f(t)$ is infinite divisible in classical sense if and only if $\varphi(-\log f(t))$ is ν-infinite divisible*. This is obviously true for strictly stable and ν-strictly stable distributions, too.

Consider some examples of ν-degenerate and ν-Gaussian distributions.

Example 1 The random variable ν_p has geometric distribution: $\mathbb{P}\{\nu_p = k\} = p(1 - p)^{k-1},\ k = 1, 2, \ldots,\ p \in (0, 1)$. In this case, ν-degenerate random variable has exponential distribution, and the ν-Gaussian distribution is the Laplace distribution (see, [5]).

Example 2 The random variable ν_p, $p \in \{1/n^2, n = 1, 2, \dots\}$ has probability generating function

$$\mathcal{P}_p(z) = 1/T_n(1/z), \quad n = 1/\sqrt{p},$$

where $T_n(z)$ is Chebyshev polynomial of the first kind ([5]). In this case, characteristic function of the ν-Gaussian distribution has form

$$f(t) = 1/\cosh(at), \quad a > 0.$$

Laplace transform of ν-degenerate distribution is

$$\varphi(s) = 1/\cosh(\sqrt{2as}), \ s > 0, \ a > 0.$$

It is clear that the both distributions are absolutely continuous and have exponential tails. The ν-Gaussian distribution is so-called hyperbolic secant distribution with density

$$p(x) = 1\Big/\left(2\cosh\left(\frac{\pi x}{2}\right)\right).$$

The density of ν-degenerate random variable is identical with that of random variable $\xi = \int_0^1 W_1^2(t)dt + \int_0^1 W_2^2(t)dt$, where $W_1(t)$, $W_2(t)$ are two independent Wiener processes (see, [8]).

4 Further Examples of ν-Gaussian Distributions

Example 3 Consider random variable ν_p with probability generating function

$$\mathcal{P}_p(z) = \frac{p^{1/m}z}{(1-(1-p)z^m)^{1/m}}, \ p \in (0, 1), \ m \in \mathbf{N},$$

which is a modification of that of negative binomial distribution. In this case, the Laplace transform of ν-degenerate (standard) distribution is

$$\varphi(s) = \frac{1}{(1+ms)^{1/m}},$$

that is Laplace transform of gamma distribution. Characteristic function of ν-Gaussian distribution is

$$f(t) = \frac{1}{(1+mat^2)^{1/m}},$$

with parameter $a > 0$. This is characteristic function of symmetrized Gamma-distribution. Let us note that negative binomial distribution does not posses ν-Gaussian distribution in algebraic sense, but has such distribution is limit sense. This is also symmetrized Gamma-distribution.

Example 4 Consider now the family of random variables $\{\nu_p,\ p \in \{1/n^2, n = 2, 3, \dots\}$ with probability generating functions

$$\mathcal{P}_p(z) = \frac{1}{(T_n(1/z^m))^{1/m}},$$

where $p = 1/n^2, n = 2, 3, \dots$ and $m \geq 1$ is an integer parameter. For corresponding standard ν-degenerate and ν-Gaussian random variables, we have Laplace transform and characteristic function

$$\varphi(s) = \frac{1}{(\cosh\sqrt{2mt})^{1/m}} \quad \text{and} \quad f(t) = \frac{1}{(\cosh at)^{1/m}},$$

where $a > 0$. Both ν-degenerate and ν-Gaussian distributions have exponential tails.

5 Toy-Model of Capital Distribution

In physics, under toy-model usually understand a model, which does not give complete description of a phenomena, but is rather simple and provides explanation of essential part of the phenomena.

Let us try to construct a toy-model for capital distribution. Assume that there is an output (business) in which we invest a unit of the capital at the initial moment $t = 0$ at the moment $t = 1$ we get a sum of capital X_1 (the nature of the r.v. X_1 depends on the nature of the output and that of the market). If the whole sum of capital remains in the business, then to the moment $t = 2$ the sum of capital becomes $X_1 \cdot X_2$, where r.v. X_2 is independent of X_1 and has the same distribution as X_1 (provided that conditions of the output and of the market are invariable). Using the same arguments further on, we find that to the moment $t = n$ the sum of capital equals to $\prod_{j=1}^{n} X_j$, and also r.v.s $X_1, \dots, X_n$ are i.i.d.

From the economical sense it is clear that $X_j > 0,\ j = 1, \dots, n$. Now assume that there can happen a change of output or of the market conditions which makes further investment of capital in the business impossible. We assume that the time till the appearance of the unfavorable event is random variable ν_p, $p = 1/\mathbb{E}\nu_p$. The sum of capital to the moment of this event equals to $\prod_{j=1}^{\nu_p} X_j$. And the mean time to the appearance of the unfavorable event is $\mathbb{E}\nu_p = 1/p$. Therefore, "mean annual sum of capital" is

$$Z_p = \Big(\prod_{j=1}^{\nu_p} X_j\Big)^p.$$

The smaller is the value of $p > 0$ the rarely is the unfavorable event. If p is small enough, we may approximate the distribution of Z_p by its limit distribution for $p \to 0$. To find this distribution it is possible to pass from X_j to $Y_j = \log X_j$, and change the product by a sum of random number ν_p of random variables Y_j.

If probability generating functions of ν_p generate a commutative semigroup, the limit distribution of the sum will coincide with ν-stable or with ν-degenerate distribution.

The most simplest case is that of geometric distribution of ν_p. In this situation, the probability of unfavorable event is the same for each time moment $t = k$. If there exists positive first moment of $Y_j = \log X_j$, then the limit distribution of random sum coincides with ν-degenerate distribution, and is Exponential distribution. This means, that limit distribution of Z_p is Pareto distribution $F(x) = 1 - x^{-1/\gamma}$ for $x > 1$, and $F(x) = 0$ for $x \leq 1$. Here $\gamma = \mathbb{E} \log X_1 > 0$. This distribution has power tail. For $\gamma \geq 1$ this distribution has infinite mean. Pareto distribution was introduced by Wilfredo Pareto to describe the capital distribution, but he used empirical study only, and had no toy-model. About hundred years ago this distribution gave a very good agreement with observed facts. Let us mention that our toy-model shows, that *such distribution of capitals may be explained just by random effects*. This is an essential argument against Elite Theory, because the definition of elite becomes not clear.

But what about capital distribution nowadays? Analysis of capitals of the first hundred billionaires (Forbes dataset) gives not bad agreement with Pareto distribution, too. However, the tail in the model is a little bit heavier than for real data (Fig. 1).

Some better agreement on the tail gives log-gamma distribution, which is log-transformation of ν-degenerate distribution for negative binomial distributed number of multipliers. In this situation, power tail of Pareto distribution is slightly corrected by logarithmic multiplier (Fig. 2).

I tried to consider capital distribution of all billionaires. Corrected Pareto distribution provides good agreement at the tail, but does not work in the central part of the distribution. This is so because the random number of summands may be considered

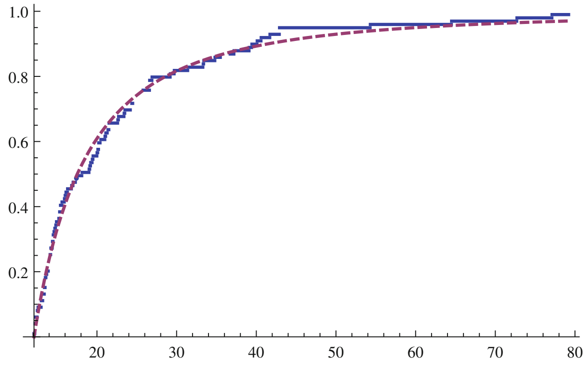


Fig. 1 Plot of Pareto distribution function (*dashed* line) versus empirical distribution of the capital of highest 100 billionaires. Forbes dataset

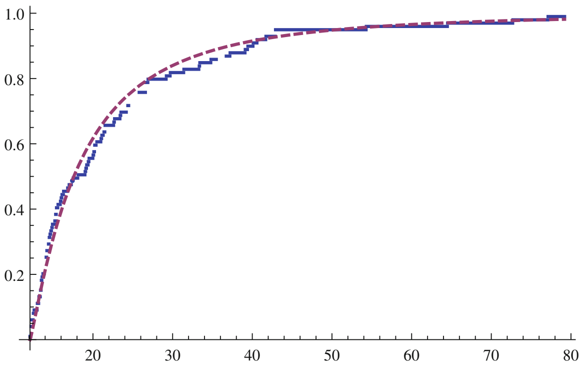


Fig. 2 Plot of log-gamma distribution function (*dashed* line) versus empirical distribution of the capital of highest 100 billionaires. Forbes dataset

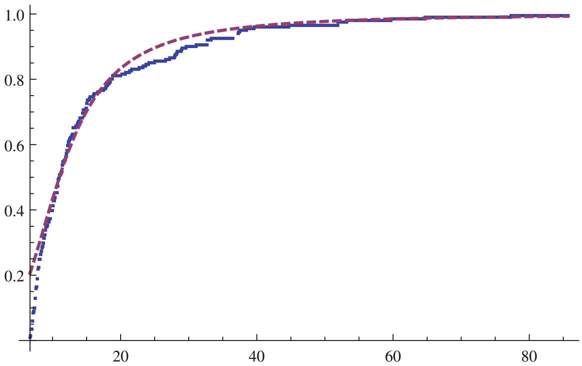


Fig. 3 Plot of ν-degenerate distribution function (*dashed* line) versus empirical distribution of the capital of all billionaires. Bloomberg dataset

now as a sum of few geometric distributions, which corresponds to "Chebyshev" probability generating function with the use of ν-degenerate distribution as possible approximation (see Example 2) (Fig. 3).

6 Few Words on the Distribution of Asset Returns

I will start with a citation from "Financial Risk and Heavy Tails" by Brendan O. Bradley and Murad S. Taqqu: "It is of great importance for those in charge of managing risk to understand how financial asset returns are distributed. Practitioners often assume for convenience that the distribution is normal. Since the 1960s, however, empirical evidence has led many to reject this assumption in favor of various heavy-tailed alternatives. In a heavy-tailed distribution the likelihood that one encounters significant deviations from the mean is much greater than in the case of the normal distribution. It is now commonly accepted that financial asset returns are, in fact, heavy-tailed. The goal of this survey is to examine how these heavy tails affect several aspects of financial portfolio theory and risk management. We describe some of the methods that one can use to deal with heavy tails and we illustrate them using the NASDAQ composite index" (see [1]).

This motivation is, in a sense, typical for papers in analysis of distributions for asset returns.

However, for asset returns distribution are used not the sums of money, but corresponding logarithms. We have seen, the distributions of logarithms of capital amounts of billionaires have *exponential* (so, not heavy!) tails. Therefore, it will be a little bit strange to expect heavy tails for asset returns distributions. In a preprint [9], there was given detailed analysis of all arguments supporting the choice of heavy tailed variant of asset returns distribution. The authors went to decision, which arguments do not really support this choice. Below I will give an analysis of the first argument of this kind.

The first argument usually arises when considering some of the time series, such as the Dow Jones Industrial Average index (say, for the interesting Period from the January 3, 2000 to December 31, 2009), daily ISE-100 Index (November 2, 1987–June 8, 2001) and many others. The observed fact is that quite a lot of data not only fall outside the 99% confidence interval on the mean, but also outside the range of $\pm 5\,\sigma$ from the average, or even $\pm 10\,\sigma$. On the assumption of this circumstance authors propose to the readers to make two conclusions.

First (and absolutely correct) conclusion consists in the fact that the observations under assumption of their independence and identical distribution are in contradiction with their normality.

The second conclusion is that the distribution of these random variables is heavy-tailed. This decision is not based on any mathematical justification. Indeed, the first thing that comes to mind is to apply the Chebyshev's inequality to the random variables with non-normal distributions. An exact inequality can be found in the book by Karlin and Studden [4]. There is also shown extremal distribution, for which the inequality becomes equality.

Let me quote the corresponding result: *Gauss Inequality. In this example we desire to determine the maximum value of*

$$\mathbb{P}\{X \in (-\infty, \mu - d] \cup [\mu + d, \infty)\},\ d > 0$$

over the class of unimodal distribution functions with mode and mean located at μ *and variance equal to* σ^2. *The solution for the case* $d^2 \geq 4\sigma^2/3$ *is given by* $4\sigma^2/9d^2$, *with rectangular part of the distribution on interval* $[\mu - 3d/2, \mu + 3d/2]$ *plus mass at* μ. *By a rectangular distribution on* $[a, b]$ *we mean a distribution* F *whose density is* $1/(b-a)$ *for* $x \in (a, b)$ *and* 0 *otherwise.*

Let us apply this result to the studied case supposing that the distribution is unimodal (in the same sense as in [4]) with finite variance (e.g., $\sigma = 1$), i.e., not a heavy-tailed distribution. We choose $\mu = 0$, $\sigma = 1$ and $d = 10\sigma = 10$. From the earlier mentioned follows

$$\mathbb{P}\{|X| > 10\} \leq \frac{1}{225},$$

and the equality is reached for the above-written distribution. It is clear that the probability of $1/225$ is not so small. Exactly in the case of a sample size of 50,000 average number of exceeds of level 10σ is more than 222 times.

Notice that samples of 50,000 are not uncommon in financial problems. Moreover, in sample such as this with $d = 40\sigma = 40$, an average number of exceeds of level 40σ will be somewhat more than 13.8. It is clear, that in this case we are not talking about the heavy tail ($\sigma = 1$!). Thus, the first argument of the appearance of heavy-tailed distributions is rejected.

Of course, used extremal distribution does not seem to be natural for describing the fluctuations of financial indexes. Especially strange looks existence of the mass at μ. Therefore, we will give an example of continuous distribution, which seems to be more natural in this case.

Let Y be a random variable with a gamma distribution with shape parameter $1/m$ and scale parameter m, where m is a positive integer. Assume Y_1, Y_2 are two independent variables and follow the distribution of Y. Define X as $Y_1 - Y_2$, this is the random variable we need. It follows symmetrized gamma distribution with the characteristic function $1/(1 + mt^2)^{1/m}$, and is, therefore, ν-Gaussian random variable for ν from Example 3, or for negative binomial ν.

It is easy to see, that this distribution has an exponential tail and hence there exist finite moments of all orders. The variance of this distribution is equal to two (i.e., $\sigma = \sqrt{2}$) for all $m > 0$. Let us give the probability$\times$10000 of deviations of the random variable X larger than 10σ from the average $m = 10$, $p = 0.589843$; $n = 30$, $p = 4.01799$; $n = 50$, $p = 6.63305$; $n = 70$, $p = 8.33146$; $n = 90$, $p = 9.44249$. Now we can say that for this distribution will be observed large deviations from mean, but the tails are exponential. We say *the distribution has pseudo-heavy tails*.

Acknowledgements Bloomberg dataset and Forbes dataset are data sets containing capitals of top 201 and 100 billionaires, respectively, and were extracted from the official web sites of Bloomberg (www.bloomberg.com) and Forbes (www.forbes.com) in year 2015. The work was partially supported by Grant GACR 16-03708S.

References

1. Bradley, B.O., Taqqu, M.S.: Financial risk and heavy tails. In: Handbook of Heavy Tailed Distributions in Finance, pp. 35–103 (2003). doi:10.1016/B978-044450896-6.50004-2
2. Gnedenko, B.V.: On some stability theorems. Lecture Notes Math. **982**, 24–31 (1983)
3. Gnegenko, B.V., Yu Korolev, V.: Random Summation. Limit Theorems and Applications. CRC Press, Boca Raton (1996)
4. Karlin, S., Studden, W.J.: Tchebycheff Systems: with Applications in Analysis and Statistics. Interscience Publishers, New York (1966)
5. Klebanov, L.B., Manija, G.M., Melamed, J.A.: A problem by V.M. Zolotarev and analogous of infinitely divisible and stable distributions in the scheme of summation of a random number of random variables. Teor. Verojatnst. i Primenen. **29**(4), 757–759 (1984)
6. Klebanov, L.B., Manija, G.M., Melamed, J.A.: ν_p-strictly stable laws and estimation of their parameters. Lecture Notes in Mathematics, vol. 1233, pp. 23–31. Springer, Berlin (1987)

7. Klebanov, L.B., Rachev, S.T.: Sums of a random number of random variables an their approximations with ν- accompanying infinitely divisible laws. Serdica Math. J. **22**(4), 471–496 (1996)
8. Klebanov, L.B., Kakosyan, A.V., Rachev, S.T., Temnov, G.: On a class of distributions stable under random summation. J. Appl. Probab. **49**, 303–318 (2012)
9. Klebanov, L.B., Volchenkova, I. Heavy Tailed Distributions in Finance: Reality or Myth? Amateurs Viewpoint, pp. 1–17 (2015). arXiv: 1507.07735v1

Smooth Estimation of Error Distribution in Nonparametric Regression Under Long Memory

Hira L. Koul and Lihong Wang

Abstract We consider the problem of estimating the error distribution function in a nonparametric regression model with long memory design and long memory errors. This paper establishes a uniform reduction principle of a smooth weighted residual empirical distribution function estimator. We also investigate consistency property of local Whittle estimator of the long memory parameter based on nonparametric residuals. The results obtained are useful in providing goodness of fit test for the marginal error distribution and in prediction under long memory.

Keywords Kernel estimation · Uniform reduction principle

1 Introduction

The nonparametric regression models with long memory errors have been discussed extensively in the recent years. See, e.g., Csörgo and Mielniczuk [1], Masry and Mielniczuk [14], Guo and Koul [6], Robinson [19], and the references therein. The main focus in these papers has been on the estimation of the regression function. It is often of interest and of practical importance to know the nature of the error distribution. The knowledge of the error distribution can improve inference about various underlying parameters in the model.

Kulik and Wichelhaus [11] studied the standard Parzen–Rosenblatt error density estimator for a nonparametric regression model with long memory errors and covariates. Lorel and Kulik [13] obtained an asymptotic expansion of the nonpara-

H.L. Koul (✉)
Department of Statistics and Probability, Michigan State University,
East Lansing, MI 48824, USA
e-mail: koul@stt.msu.edu

L. Wang
Department of Mathematics, Nanjing University, Nanjing 210093,
People's Republic of China
e-mail: lhwang@nju.edu.cn

J. Antoch et al. (eds.), *Analytical Methods in Statistics*, Springer Proceedings
in Mathematics & Statistics 193, DOI 10.1007/978-3-319-51313-3_5

metric residual empirical process under broad conditions on the covariate process in nonparametric regression models with long memory moving average errors.

In this paper, we investigate the smooth estimators of the error distribution function (d.f.) in the same set up as in Kulik and Wichelhaus [11]. The proposed estimator is a weighted kernel type estimator based on nonparametric residuals where regression function is estimated by a modified Nadaraya–Watson estimator.

More precisely, consider a random process defined as a collection of random vectors $(X_i, Y_i) \in \mathbb{R} \times \mathbb{R}$, where X_i denotes the ith covariate and Y_i the corresponding response. We assume that the process $\{X_i, Y_i; i \in \mathbb{Z}\}$, $\mathbb{Z} := \{0, \pm 1, \pm 2, \cdots\}$ is stationary and $\mathrm{E}|Y_0| < \infty$. With $\mu(x) = \mathrm{E}(Y_0|X_0 = x)$, the data consists of (X_i, Y_i), $1 \le i \le n$, obeying the model

$$Y_i = \mu(X_i) + \varepsilon_i, \qquad \varepsilon_i = \sum_{j \le i} a_{i-j}\zeta_j, \quad i \in \mathbb{Z}, \tag{1}$$

where ζ_j, $j \in \mathbb{Z}$, are i.i.d. standardized r.v.'s.

Moreover, a_j, $j \in \mathbb{Z}$, are nonrandom weights such that for some $d \in (0, 1/2)$ and $0 < c < \infty$,

$$a_j \sim c j^{d-1}, \quad j \to \infty.$$

By Proposition 3.2.1 of Giraitis, Koul and Surgailis [5],

$$\gamma(i) = \mathrm{Cov}(\varepsilon_0, \varepsilon_i) \sim c^2 B(d, 1-2d)|i|^{2d-1}, \quad |i| \to \infty, \tag{2}$$

where $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$, $\alpha > 0$, $\beta > 0$. This implies that the error process $\{\varepsilon_i,\ i \in \mathbb{Z}\}$ has long memory in the covariance sense.

We additionally assume that the covariate process $\{X_i,\ i \in \mathbb{Z}\}$ satisfy

$$X_i = \mu_X + \sum_{j \le i} b_{i-j}\eta_j, \quad i \in \mathbb{Z}, \quad \mu_X \in \mathbb{R}, \tag{3}$$

where η_j, $j \in \mathbb{Z}$, are i.i.d. standardized r.v.'s, independent of $\{\zeta_i,\ i \in \mathbb{Z}\}$, and $b_j \sim c_X j^{d_X - 1}$, $j \in \mathbb{Z}$, as $j \to \infty$ for some $d_X \in (0, 1/2)$ and $0 < c_X < \infty$. Thus, the process $\{X_i,\ i \in \mathbb{Z}\}$ is independent of the process $\{\varepsilon_i, i \in \mathbb{Z}\}$, and has long memory because for some $0 < C_X < \infty$,

$$\gamma_X(i) = \mathrm{Cov}(X_0, X_i) \sim C_X |i|^{2d_X - 1}, \quad |i| \to \infty.$$

Additional needed assumptions will be described in the next section.

Now, let f and F denote the common marginal density function and d.f. of ε_0, respectively. The problem of interest is to provide a goodness-of-fit test for F. Koul and Surgailis [8, 9] observed that in the one sample location model, and more generally in the parametric multiple linear regression models with non-zero intercept and long memory errors, the first order asymptotic behavior of the residual empirical process based on the least square residuals is degenerate. Kulik and Wichelhaus [11] also showed this degeneracy phenomenon for the weak limit of the kernel error density estimator. Because of this first order degeneracy, these estimators can not be used to develop useful goodness-of-fit tests for F or f in these models. To overcome this deficiency, Koul, Mimoto, and Surgailis [7] proposed a modified estimator of the location parameter so that the corresponding residual empirical process in the one sample location model is not the first order degenerate asymptotically, and investigated some tests for fitting an error d.f. based on this estimator. In this paper we extend this methodology to long memory nonparametric regression set up.

To proceed further, we need to obtain nonparametric residuals. Let $\tilde{K}$ be a kernel density function on $\mathbb{R}$, h_n be a bandwidth sequence, and let $\tilde{K}_{h_n}(\cdot) = h_n^{-1}\tilde{K}(\cdot/h_n)$. Let ϕ be a piece-wise continuously differentiable function on [0, 1] and define

$$\phi_{ni} = n\int_{((i-1)/n, i/n]} \phi(u)du, \quad 1 \le i \le n, \tag{4}$$
$$f_{n,X}(x) = \frac{1}{n}\sum_{i=1}^{n} \tilde{K}_{h_n}(x - X_i), \quad \hat{\mu}_n(x) = \frac{1}{n}\sum_{i=1}^{n} Y_i(1+\phi_{ni})\tilde{K}_{h_n}(x - X_i)/f_{n,X}(x), \quad x \in \mathbb{R},$$

Note that $f_{n,X}$ and $\hat{\mu}_n$ are the kernel estimators of the density f_X of X_0 and the regression function μ, respectively. Use $\hat{\mu}_n$ to define the residuals

$$\hat{\varepsilon}_i := Y_i - \hat{\mu}_n(X_i), \quad 1 \le i \le n.$$

The classical estimator of F is the empirical d.f. However, for continuous F it seems more appropriate to use a smooth estimator. As proposed in Fernholz [3], Liu and Yang [12] and Wang et al. [20], we use a kernel type estimator to estimate the function F. In addition, as noted in Müller, Schick and Wefelmeyer [15], since the performance of the estimator $\hat{\mu}_n(x)$ will be poor for large values of x, we shall use only the residuals $\hat{\varepsilon}_i$ for which X_i falls into an interval $[\tau_{1n}, \tau_{2n}]$, where $\tau_{1n} < 0 < \tau_{2n}$, and $-\tau_{1n}$ and τ_{2n} tend to infinity slowly.

Let $w(\cdot)$ be a continuous weight function that vanishes off $[\tau_{1n}, \tau_{2n}]$, is 1 on $[\tau_{1n} + r, \tau_{2n} - r]$ for some fixed small positive r and is linear on the intervals $[\tau_{1n}, \tau_{1n} + r]$ and $[\tau_{2n} - r, \tau_{2n}]$. Define the random weights

$$w_i = \frac{w(X_i)}{\sum_{j=1}^{n} w(X_j)}, \quad 1 \le i \le n. \tag{5}$$

Let K be another kernel density function and b_n be another bandwidth sequence tending to zero, and define

$$\hat{F}(x) = \sum_{i=1}^{n} w_i \int_{-\infty}^{x} K_{b_n}(u - \hat{\varepsilon}_i)du, \quad x \in \mathbb{R}. \tag{6}$$

This estimator can be used to fit a known error d.f. This will be elaborated further on in the next section. It may be also used to construct prediction intervals as follows. Given an observation X_* of the covariate, the predication of the corresponding response Y_* can be defined as $\hat{Y}_* = \hat{\mu}_n(X_*)$. Then, using the estimator $\hat{F}(x)$, we can construct a prediction interval $[\hat{Y}_* + \hat{F}^{-1}(\alpha_1), \hat{Y}_* + \hat{F}^{-1}(\alpha_2)]$ for Y_*, with the asymptotic confidence level $\alpha_2 - \alpha_1$, $0 < \alpha_1 < \alpha_2 < 1$, provided that we know the asymptotic distribution of the estimator $\hat{F}$. The asymptotic properties of $\hat{F}(x)$ are described in Sect. 2 below, along with the needed assumptions.

Throughout the paper, all limits are taken as $n \to \infty$, unless specified otherwise, $\to_D$ denotes convergence in distribution, $\to_p$ denotes convergence in probability, $\|\cdot\|$ denote the supremum norm, and Z stands for a standard normal random variable.

2 Main Results

In this section we describe the first order large sample behavior of the above $\hat{F}$. To begin with, we state the needed assumptions for the kernels K, $\tilde{K}$, the density f_X, the function μ, and the bandwidths h_n and b_n. Let $\mathbf{i} := \sqrt{-1}$.

Assumption (K): K is a symmetric, bounded, and differentiable density with bounded derivative K' and $\int_{-\infty}^{\infty} u^2 K(u)du < \infty$.

Assumption ($\tilde{\text{K}}$): $\tilde{K}$ is a symmetric differentiable density with derivative $\tilde{K}'$ satisfying $\int_{-\infty}^{\infty} |\tilde{K}'(u)|du < \infty$ and $\int_{-\infty}^{\infty} u^2 \tilde{K}(u)du < \infty$.

Assumption (A): The distribution of ζ_j in (1) satisfies the following two conditions: there exists constants $C, \delta > 0$ such that $|\mathrm{E}e^{\mathbf{i}t\zeta_j}| \le C(1+|t|)^{-\delta}$, $t \in \mathbb{R}$ and $\mathrm{E}|\zeta_j|^3 < \infty$.

Assumption (B):

(B1) The function μ is twice differentiable with bounded integrable derivatives, and $\mathrm{E}\mu^2(X_0) < \infty$.

(B2) The density f_X is positive on $\mathbb{R}$ with derivative f_X' having finite Fisher information for location, i.e., $\int_{-\infty}^{\infty} |f_X'(x)|^2/f_X(x)dx < \infty$.

(B3) The distribution of η_j in (3) satisfies the following two conditions: there exists constants $C_X, \delta_X > 0$ such that $|\mathrm{E}e^{\mathbf{i}t\eta_j}| \le C_X(1+|t|)^{-\delta_X}$, $t \in \mathbb{R}$ and $\mathrm{E}|\eta_j|^3 < \infty$.

(B4) The interval $[\tau_{1n}, \tau_{2n}]$ is such that $-\tau_{1n} = o(n^\delta)$ and $\tau_{2n} = o(n^\delta)$ tend to infinity slowly enough so that $(\log n) \cdot \inf_{x \in [\tau_{1n}, \tau_{2n}]} f_X(x) \ge M$, for some $0 < M < \infty$ and any $\delta > 0$.

(B5) $h_n^{1/2} b_n^{-1} \to 0$, $n^{-d} h_n^{-1/2} b_n^{-1} \log n \to 0$, $n^{1/2-d} h_n^2 b_n^{-1} \log n \to 0$, $n^{d_X - d} b_n^{-1} \log n \to 0$, $n^{d_X - 1/2} h_n^{-1} b_n^{-1} \log n \to 0$, $n^{1/2-d} b_n^2 \to 0$, and $0 < d_X < d < 1/2$.

Remark 1 Assumption (B5) actually implies the following assumption:

(B5′) $n^{-d}h_n^{-1/2}\log n \to 0$, $n^{1/2-d}h_n^2\log n \to 0$, $n^{d_X-1/2}h_n^{-1}\log n \to 0$, and $0 < d_X < d < 1/2$.

As an example, consider $h_n \sim n^{-a}$, $a > 0$. Then assumption (B5′) will be satisfied as long as $1/4 - d/2 < a < \min(2d, 1/2 - d_X)$ for $1/10 < d < 1/2$ and $d_X < d$. In addition, if $b_n \sim n^{-\rho}$, $\rho > 0$, then Assumption (B5) holds for $1/4 - d/2 < \rho < \min(a/2, d - a/2, d + 2a - 1/2, d - d_X, 1/2 - d_X - a)$. For example, if $d = 0.4$, $d_X = 0.2$, then $0.05 < a < 0.3$. Let $a = 0.2$, then $0.05 < \rho < 0.1$.

The Assumptions (K) and ($\tilde{\text{K}}$) are the usual standard conditions for the kernel type estimation. Assumption (K) implies that $\int_{-\infty}^{\infty}|u|^{1/2}K(u)du < \infty$ and Assumption ($\tilde{\text{K}}$) implies that $\int_{-\infty}^{\infty}|u|^j\tilde{K}^2(u)du < \infty$, $j = 0, 1, 2$. Examples of the kernels satisfying Assumptions (K) and ($\tilde{\text{K}}$) are the Gaussian kernel and the uniform kernel vanishing off $(-1, 1)$.

Assumptions similar to (A) and (B3) are imposed for the first time in Giraitis, Koul, and Surgailis [4] for studying the empirical processes of long memory sequences. Under Assumptions (A) and (B3), Koul and Surgailis [8] (see also Lemma 10.2.4 of Giraitis et al. [5]) showed that the densities f and f_X are bounded and continuously infinitely differentiable, having bounded derivatives of all orders. Perhaps it is worth mentioning that Gaussian distribution satisfies Assumptions (A), (B2), and (B3).

Furthermore, from Lemma 2 of Giraitis et al. [4], Assumption (A) implies that as $|i| \to \infty$, $|j| \to \infty$ and $|i - j| \to \infty$,

$$
\begin{aligned}
f_{0,i}(x, y) - f(x)f(y) &= \gamma(i)f'(x)f'(y) + O(|i|^{2d-1-\alpha}),\\
f_{0,i,j}(x, y, z) - f(x)f_{0,j-i}(y, z) &= \gamma(i)f'(x)\frac{\partial f_{0,j-i}(y, z)}{\partial y} + \gamma(j)f'(x)\frac{\partial f_{0,j-i}(y, z)}{\partial z}\\
&\quad + O(|i|^{2d-1-\alpha}) + O(|j|^{2d-1-\alpha}),
\end{aligned}
$$

uniformly in $x, y, z \in \mathbb{R}$, for any small positive number α. Here, $f_{0,i}$ and $f_{0,i,j}$ are the joint densities of $(\varepsilon_0, \varepsilon_i)$ and $(\varepsilon_0, \varepsilon_i, \varepsilon_j)$, respectively, and $\partial f_{0,j-i}(y, z)/\partial y$ and $\partial f_{0,j-i}(y, z)/\partial z$ are the partial derivatives of $f_{0,j-i}(y, z)$.

In addition, we shall prove in the Appendix that Assumption (B3) implies that as $|i| \to \infty$, $|j| \to \infty$ and $|i - j| \to \infty$,

$$
\begin{aligned}
(B3')\quad & f_{X,0,i}(x, y) - f_X(x)f_X(y) - \gamma_X(i)f_X'(x)f_X'(y) = O(|i|^{2d_X-1-\alpha}),\\
& f_{X,0,i,j}(x, y, z) - f_X(x)f_{X,0,j-i}(y, z) - \gamma_X(i)f_X'(x)\frac{\partial f_{X,0,j-i}(y, z)}{\partial y}\\
&\qquad - \gamma_X(j)f_X'(x)\frac{\partial f_{X,0,j-i}(y, z)}{\partial z} = O(|i|^{2d_X-1-\alpha}) + O(|j|^{2d_X-1-\alpha}),
\end{aligned}
$$

uniformly $x, y, z \in \mathbb{R}$, where α is a small positive number, $f_{X,0,i}$ and $f_{X,0,i,j}$ are the joint densities of (X_0, X_i) and (X_0, X_i, X_j), respectively, and $\partial f_{X,0,j-i}(y, z)/\partial y$ and $\partial f_{X,0,j-i}(y, z)/\partial z$ are the partial derivatives of $f_{X,0,j-i}(y, z)$. This property actually defines the long memory behavior in distribution for the error and covariate processes. The proof of the claim (B3) implies (B3′) is given in the Appendix.

We now state some needed preliminaries. With ϕ_{ni} as in (4), let

$$\bar{\varepsilon}_n = \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i, \quad \bar{W}_n = \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i\phi_{ni}. \tag{7}$$

Let $c_0^2 = c^2 B(d, 1-2d)/(d(1+2d))$ and Z be a standard normal r.v., and

$$c^2(\phi) = c^2 B(d, 1-2d)\int_{[0,1]^2}\phi(u)\phi(v)|u-v|^{2d-1}dudv.$$

From Lemma 2.1 of Koul et al. [7],

$$n^{1/2-d}\bar{\varepsilon}_n \to_D c_0 Z, \qquad n^{1-2d}\mathrm{E}\bar{\varepsilon}_n^2 = c_0^2 + o(1), \tag{8}$$
$$n^{1/2-d}\bar{W}_n \to_D W = c(\phi)Z. \tag{9}$$

Theorem 1 below establishes the uniform reduction principle, and hence the asymptotic distribution of the proposed kernel estimator $\hat{F}(x)$.

Theorem 1 *Suppose that the Assumptions (K), ($\tilde{K}$), (A), and (B) hold. Let $\phi(u)$, $u \in [0, 1]$ be a piecewise continuously differentiable function satisfying $\bar{\phi} = 0$. Then*

$$\sup_{x\in\mathbb{R}}|\hat{F}(x) - F(x) - \bar{W}_n f(x)| = o_p(n^{d-1/2}). \tag{10}$$

Consequently, $n^{1/2-d}\big(\hat{F}(\cdot) - F(\cdot)\big)$ converges weakly, in Skorokhod space $D(\mathbb{R})$ and uniform metric, to the degenerate Gaussian process $f(\cdot)c(\phi)Z$.

An immediate application of the above theorem is to the goodness-of-fit testing. Let F_0 be a known d.f. satisfying the conditions of the above theorem, with f_0 denoting its density. Consider the problem of testing $H_0 : F = F_0$, versus the alternative that H_0 is not true. Let $\hat{d}$ be an estimator of d such that $(\log n)(\hat{d} - d) \to_p 0$, and let $\hat{c}(\phi)$ be a consistent estimator of $c(\phi)$, under H_0. Let

$$\mathscr{D}_n := n^{1/2-\hat{d}}\|\hat{F} - F_0\|\big/[\hat{c}(\phi)\,\|f_0\|].$$

By (9) and (10), under H_0, $\mathscr{D}_n \to_D |Z|$. Hence, the test that rejects H_0 whenever $\mathscr{D}_n > z_{\alpha/2}$ will have the asymptotic size α, where $z_{\alpha/2}$ is $(1-\alpha/2)100$th percentile of the standard normal distribution. Section 3 below discusses some estimators of d and $c(\phi)$, which satisfy the said consistency conditions.

Now we turn to the proof of Theorem 1. The claim about the weak convergence readily follows from (9) and (10). The proof of (10) is facilitated by the following several lemmas. Before stating the lemmas, we first describe some properties of the random weights w_i of (5), which will be used in the proofs later. The indices in all sums vary from 1 to n, unless mentioned otherwise.

Let $n\bar{w} = \sum_{i=1}^{n} w(X_i)$. By definition, $0 \le w(x) \le 1$ and $|w(X_0) - 1|^2 \le |w(X_0) - 1|$. Hence, stationarity of X_i's implies

$$
\begin{aligned}
E|w(X_0)-1| &\le \Big|\int_{\tau_{1n}+r}^{\tau_{2n}-r} f_X(x)dx - 1\Big| + \int_{\tau_{1n}}^{\tau_{1n}+r} f_X(x)dx + \int_{\tau_{2n}-r}^{\tau_{2n}} f_X(x)dx \\
&\to 0, \\
E|w(X_0)-1|^2 &\to 0, \\
E|\bar{w}-1|^2 &\le E|\bar{w}-1| \le E|w(X_0)-1| \to 0. \qquad (11)
\end{aligned}
$$

These facts in turn imply

$$
\bar{w} = 1 + o_p(1), \quad \max_{1\le i\le n} |w_i - n^{-1}w(X_i)| = o_p(1), \max_{1\le i\le n} |w_i| = O_p(n^{-1}). \qquad (12)
$$

$$
\begin{aligned}
\max_{1\le i\le n} \mathrm{E}\big|\bar{w}(n^{-1}-w_i)\big|^2 &= n^{-2}\max_{1\le i\le n} \mathrm{E}|\bar{w}-w(X_i)|^2 \qquad (13)\\
&\le 2n^{-2}\mathrm{E}|\bar{w}-1|^2 + 2n^{-2}\mathrm{E}|w(X_0)-1|^2 = o(n^{-2}).
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
\int_{-\infty}^{\infty} w(x)dx &= r^{-1}\int_{\tau_{1n}}^{\tau_{1n}+r}(x-\tau_{1n})dx + \int_{\tau_{1n}+r}^{\tau_{2n}-r} dx + r^{-1}\int_{\tau_{2n}-r}^{\tau_{2n}}(\tau_{2n}-x)dx \\
&= (\tau_{2n}-\tau_{1n}) - r.
\end{aligned}
$$

In view of this fact, (B3′) and assumption (B4), we also have

$$
\begin{aligned}
&\sum_{t,s=1}^{n} \big|\mathrm{Cov}(w(X_t), w(X_s))\mathrm{E}(\varepsilon_t\varepsilon_s)\big| \qquad (14)\\
&= n\mathrm{Var}(w(X_0))E(\varepsilon_0^2) \\
&\quad + \sum_{s\ne t} |\gamma(t-s)|\Big|\int_{-\infty}^{\infty} w(x)w(y)\big[f_{X,s,t}(x,y) - f_X(x)f_X(y)\big]dxdy\Big| \\
&= O(n) + \sum_{s\ne t} O(|t-s|^{2d+2d_X-2})\Big(\int_{-\infty}^{\infty}\big|f_X'(y)\big|dy\Big)^2 \\
&\qquad + \sum_{s\ne t} O(|t-s|^{2d+2d_X-2-\alpha}(\tau_{2n}-\tau_{1n})^2) \\
&= O(n) + O(n^{2d+2d_X}),
\end{aligned}
$$

upon choosing $\delta = \alpha/2$ in (B4). The finiteness of the integral $\int_{-\infty}^{\infty}|f_X'(y)|dy$ is guaranteed by (B2).

We are now ready to state and prove the first lemma.

Lemma 1 *Under Assumptions ($\tilde{K}$) and (B3),*

$$
\sup_{x\in\mathbb{R}} |f_{n,X}(x) - f_X(x)| = O_p(\max\{n^{d_X-1/2}h_n^{-1},\ h_n\}).
$$

Proof Let $U_i := X_i - \mu_X$, $\bar{U}_n = n^{-1}\sum_{i=1}^n U_i$, $\mathscr{G}$ denote the d.f. of U_0, g its density, and let $\mathscr{G}_n(y) := n^{-1}\sum_{i=1}^n I(U_i \le y)$. From (10.3.31), (10.3.32) and Theorem 10.2.3 of Giraitis et al. [5], we obtain that under (3) and assumption (B3),

$$|\bar{U}_n| = O_p(n^{d_X - 1/2}), \quad \sup_{y\in\mathbb{R}} \Big|\mathscr{G}_n(y) - \mathscr{G}(y) + g(y)\bar{U}_n\Big| = o_p(n^{d_X-1/2}). \tag{15}$$

Now write

$$\begin{aligned} f_{n,X}(x) &= \frac{1}{nh_n}\sum_{i=1}^n \tilde{K}(\frac{x - X_i}{h_n}) = \frac{1}{nh_n}\sum_{i=1}^n \tilde{K}(\frac{x - \mu_X - (X_i - \mu_X)}{h_n}) \\ &= \frac{1}{h_n}\int_{-\infty}^{\infty} \tilde{K}\big(\frac{x - \mu_X - y}{h_n}\big) d\mathscr{G}_n(y), \\ \mathrm{E} f_{n,X}(x) &= \frac{1}{h_n}\int_{-\infty}^{\infty} \tilde{K}\big(\frac{x - \mu_X - y}{h_n}\big) d\mathscr{G}(y). \end{aligned}$$

Hence, integration by parts and a change of variables yield

$$\begin{aligned} & f_{n,X}(x) - \mathrm{E} f_{n,X}(x) \\ &= \frac{1}{h_n}\int_{-\infty}^{\infty} \tilde{K}\big(\frac{x - \mu_X - y}{h_n}\big) d(\mathscr{G}_n(y) - \mathscr{G}(y)) \\ &= \frac{1}{h_n}\int_{-\infty}^{\infty} [\mathscr{G}_n(x - \mu_X - h_n z) - \mathscr{G}(x - \mu_X - h_n z)]\tilde{K}'(z)dz \\ &= \frac{1}{h_n}\int_{-\infty}^{\infty} [\mathscr{G}_n(x - \mu_X - h_n z) - \mathscr{G}(x - \mu_X - h_n z) + g(x - \mu_X - h_n z)\bar{U}_n]\tilde{K}'(z)dz \\ &\qquad - \frac{1}{h_n}\int_{-\infty}^{\infty} g(x - \mu_X - h_n z)\tilde{K}'(z)dz\ \bar{U}_n. \end{aligned}$$

Thus, by (15) and assumptions ($\tilde{K}$) and (B3), which implies g is bounded,

$$\sup_{x\in\mathbb{R}} \big| f_{n,X}(x) - \mathrm{E} f_{n,X}(x)\big| = o_p(h_n^{-1} n^{d_X-1/2}) + O_p(h_n^{-1} n^{d_X-1/2}).$$

The proof of the lemma is completed upon noting that the boundedness of f_X and its derivative (see Remark 1) implies $\sup_{x\in\mathbb{R}} |\mathrm{E} f_{n,X}(x) - f_X(x)| = O(h_n)$.

To state the next result, recall (5) and (7) and let

$$\begin{aligned} &\hat{Z}_i = \varepsilon_i - \hat{\varepsilon}_i, \ 1 \le i \le n, \quad Z_n = \bar{\varepsilon}_n + \bar{W}_n, \\ &\gamma_n = \max(h_n^2 \log n,\ n^{-1/2}h_n^{-1/2}\log n,\ n^{d_X-1/2}\log n,\ n^{d-1/2}h_n^{1/2}), \\ &\xi_n = \max(\gamma_n,\ n^{d+d_X-1}h_n^{-1}\log n). \end{aligned} \tag{16}$$

We are now ready to state and prove

Lemma 2 *Assume ($\tilde{K}$), (B1)–(B4), (B5′), and that ϕ is a piecewise continuously differentiable function satisfying $\bar{\phi} = 0$. Then the following holds.*

$$\sum_{i=1}^{n}(\hat{Z}_i - Z_n)^2 w_i = O_p(\xi_n^2).$$

The proof of this lemma is facilitated by the following lemma.

Lemma 3 *Under the conditions of Lemma 2, the following holds.*

$$\sum_{i=1}^{n}\Big(\frac{f_{n,X}(X_i)}{f_X(X_i)}\hat{Z}_i - Z_n\Big)^2 w_i = O_p(\gamma_n^2). \tag{17}$$

Proof Let $\tilde{K}_j(X_i) = \tilde{K}((X_i - X_j)/h_n)$. From (1) and (4), we obtain

$$\begin{aligned} \frac{f_{n,X}(X_i)}{f_X(X_i)}\hat{Z}_i - Z_n &= \frac{f_{n,X}(X_i)}{f_X(X_i)}(\hat{\mu}_n(X_i) - \mu(X_i)) - Z_n \\ &= \frac{1}{nh_n}\sum_{j=1}^{n}\varepsilon_j(1+\phi_{nj})\frac{\tilde{K}_j(X_i)}{f_X(X_i)} - Z_n \\ &\quad + \frac{1}{f_X(X_i)}\frac{1}{nh_n}\sum_{j=1}^{n}(\mu(X_j) - \mu(X_i))\tilde{K}_j(X_i) \\ &\quad + \frac{1}{f_X(X_i)}\frac{1}{nh_n}\sum_{j=1}^{n}\mu(X_j)\phi_{nj}\tilde{K}_j(X_i). \end{aligned} \tag{18}$$

Let

$$H_{ni} = \frac{1}{nh_n}\sum_{j=1}^{n}(\mu(X_j) - \mu(X_i))\tilde{K}_j(X_i) = \frac{1}{nh_n}\sum_{j\neq i}(\mu(X_j) - \mu(X_i))\tilde{K}_j(X_i).$$

Then

$$\begin{aligned} \mathrm{E}\big(H_{ni}^2\big) &= \frac{1}{n^2 h_n}\sum_{j\neq i}\int_{-\infty}^{\infty}[\mu(x - h_n u) - \mu(x)]^2\tilde{K}^2(u) f_{X,i,j}(x, x - h_n u)dudx \\ &\quad + \frac{1}{n^2}\sum_{j\neq k;\, j,k\neq i}\int_{-\infty}^{\infty}(\mu(x - h_n u) - \mu(x))(\mu(x - h_n v) - \mu(x))\tilde{K}(u)\tilde{K}(v) \\ &\qquad\qquad \times f_{X,i,j,k}(x, x - h_n u, x - h_n v)dxdudv \\ &=: I_{1i} + I_{2i}, \qquad \text{say.} \end{aligned}$$

By (B1) and (B3) (or (B3′)), the fact $d_X < 1/2$, for some large enough N,

$$\begin{aligned}
&\sum_{i=1}^{n} I_{1i} \qquad (19)\\
&\leq Cn^{-2}h_n \sum_{i=1}^{n}\sum_{j\neq i}\int_{-\infty}^{\infty} |u|^2\tilde{K}^2(u)\mu'^2(\xi) f_{X,i,j}(x, x-h_n u)dudx\\
&= Cn^{-2}h_n \sum_{i=1}^{n}\sum_{j\neq i}\int_{-\infty}^{\infty} |u|^2\tilde{K}^2(u) f_X(x) f_X(x-h_n u)dxdu\\
&\quad + Cn^{-2}h_n \sum_{i=1}^{n}\sum_{j\neq i}\int_{-\infty}^{\infty} |u|^2\tilde{K}^2(u)\mu'^2(\xi)\\
&\qquad\qquad \times\Big[f_{X,i,j}(x, x-h_n u) - f_X(x) f_X(x-h_n u)\Big]dxdu\\
&= O(h_n) + O(n^{-1}h_n) + Cn^{-2}h_n \sum_{|j-i|>N}\Big(|j-i|^{2d_X-1} + o(|j-i|^{2d_X-1})\Big)\\
&= O(h_n) + O(n^{2d_X-1}h_n) = O(h_n),
\end{aligned}$$

where ξ is between x and $x - h_n u$.

To deal with I_{2i} terms, note that by (B1),

$$\begin{aligned}
&(\mu(x-h_n u)-\mu(x))(\mu(x-h_n v)-\mu(x)) f_X(x) f_X(x-h_n u) f_X(x-h_n v) \qquad (20)\\
&\quad = h_n^2 uv\mu'(x)^2 f_X(x) f_X(x-h_n u) f_X(x-h_n v)\\
&\qquad + \frac{1}{2}h_n^3 uv^2\mu'(x)\mu''(\xi_v) f_X(x) f_X(x-h_n u) f_X(x-h_n v)\\
&\qquad + \frac{1}{2}h_n^3 u^2 v\mu'(x)\mu''(\xi_u) f_X(x) f_X(x-h_n u) f_X(x-h_n v)\\
&\qquad + \frac{1}{4}h_n^4 u^2v^2\mu''(\xi_u)\mu''(\xi_v) f_X(x) f_X(x-h_n u) f_X(x-h_n v),
\end{aligned}$$

where ξ_u, ξ_v are between x and $x - h_n u$ and x and $x - h_n v$, respectively. Moreover,

$$\begin{aligned}
&h_n^2 uv\mu'(x)^2 f_X(x) f_X(x-h_n u) f_X(x-h_n v)\\
&\quad = h_n^2 uv\mu'(x)^2 f_X^3(x) - h_n^3 uv^2\mu'(x)^2 f_X^2(x) f_X'(x)\\
&\qquad - h_n^3 u^2 v\mu'(x)^2 f_X^2(x) f_X'(x) + O(h_n^4)u^2v^2 f_X(x).
\end{aligned}$$

Similar equations are valid for the second and third terms on the right side of (20). Therefore, by ($\tilde{K}$), since $\int u\tilde{K}(u)du = 0$, we obtain

$$I_{2i} \leq \frac{Ch_n^4}{n^2}\sum_{j\neq k;\, j,k\neq i}\int_{-\infty}^{\infty} |uv|^2\tilde{K}(u)\tilde{K}(v) f_X(x)dxdudv$$

$$+\frac{Ch_n^2}{n^2}\sum_{j\neq k;\,j,k\neq i}\int_{-\infty}^{\infty}|uv|\tilde{K}(u)\tilde{K}(v)|\mu'(x)|\Big|f_{X,i,j,k}(x,x-h_nu,x-h_nv)$$
$$-f_X(x)f_{X,j,k}(x-h_nu,x-h_nv)\Big|dxdudv$$
$$+\frac{Ch_n^2}{n^2}\sum_{j\neq k;\,j,k\neq i}\int_{-\infty}^{\infty}|uv|\tilde{K}(u)\tilde{K}(v)|\mu'(x)|f_X(x)$$
$$\times\Big|f_{X,j,k}(x-h_nu,x-h_nv)-f_X(x-h_nu)f_X(x-h_nv)\Big|dxdudv$$
$$=: R_{1i}+R_{2i}+R_{3i},\qquad \text{say.}$$

Clearly,

$$\sum_{i=1}^{n}R_{1i}=O(nh_n^4).$$

By Assumption (B3) and arguing as in the proof of (19),

$$\sum_{i=1}^{n}R_{2i}\le Ch_n^2+C'n^{-1}h_n^2\Big\{\sum_{|j-i|>N}\Big(|j-i|^{2d_X-1}+o(|j-i|^{2d_X-1})\Big)$$
$$+\sum_{|k-i|>N}\Big(|k-i|^{2d_X-1}+o(|k-i|^{2d_X-1})\Big)\Big\}$$
$$=O(h_n^2)+O(n^{2d_X}h_n^2).$$

Similarly,

$$\sum_{i=1}^{n}R_{3i}=O(h_n^2)+O(n^{2d_X}h_n^2).$$

Combine these bounds with the fact $n^{2d_X}h_n^2/nh_n^4=n^{2d_X-1}h_n^{-2}\to 0$, to obtain

$$\sum_{i=1}^{n}I_{2i}=O(nh_n^4).$$

This bound together with (19) yields

$$\mathrm{E}\Big(\sum_{i=1}^{n}H_{ni}^2\Big)=O(\max\{h_n,nh_n^4\}).$$

Then it follows from (12) and Assumption (B4) that

$$\sum_{i=1}^{n} \frac{w_i}{f_X^2(X_i)} \Big(\frac{1}{nh_n} \sum_{j=1}^{n} (\mu(X_j) - \mu(X_i)) \tilde{K}_j(X_i)\Big)^2 \tag{21}$$
$$\leq \max_{1 \leq i \leq n} \frac{w_i}{f_X^2(X_i)} \sum_{i=1}^{n} H_{ni}^2 = O_p(\max\{n^{-1} h_n \log^2 n,\ h_n^4 \log^2 n\}).$$

Next, consider the third term in the right hand side of (18). Note that

$$\mathrm{E}\Big(\frac{1}{nh_n} \sum_{j=1}^{n} \mu(X_j)\phi_{nj} \tilde{K}_j(X_i)\Big)^2$$
$$= \frac{\phi_{ni}^2 \tilde{K}^2(0)}{n^2 h_n^2} \mathrm{E}\mu^2(X_i) + \frac{1}{n^2 h_n^2} \sum_{j \neq i} \phi_{nj}^2 \mathrm{E}\Big(\mu(X_j) \tilde{K}_j(X_i)\Big)^2$$
$$+ \frac{2\phi_{ni} \tilde{K}(0)}{n^2 h_n^2} \sum_{j \neq i} \phi_{nj} \mathrm{E}\Big(\mu(X_i)\mu(X_j)\tilde{K}_j(X_i)\Big)$$
$$+ \frac{1}{n^2 h_n^2} \sum_{j \neq k; j,k \neq i} \phi_{nj}\phi_{nk} \mathrm{E}\Big(\mu(X_j)\tilde{K}_j(X_i)\mu(X_k)\tilde{K}_k(X_i)\Big)$$
$$= Q_{0i} + Q_{1i} + Q_{2i} + Q_{3i}, \qquad \text{say.}$$

Standard kernel arguments yield that

$$\sum_{i=1}^{n} Q_{0i} = O(n^{-1} h_n^{-2}), \qquad \sum_{i=1}^{n} Q_{2i} = O(h_n^{-1}),$$
$$\sum_{i=1}^{n} Q_{1i} = \frac{1}{n^2 h_n^2} \sum_{i=1}^{n} \sum_{j \neq i} \phi_{nj}^2 \int_{-\infty}^{\infty} \mu^2(y) \tilde{K}^2\Big(\frac{x-y}{h_n}\Big) f_{X,i,j}(x,y)\,dx\,dy = O(h_n^{-1}).$$

For the sake of brevity, let $dK(u,v) := \tilde{K}(u)\tilde{K}(v)dudv$. Then, by Assumption (B3) and $\bar{\phi} = 0$, we obtain

$$Q_{3i} = \frac{1}{n^2} \sum_{j \neq k; j,k \neq i} \phi_{nj}\phi_{nk} \int_{-\infty}^{\infty} \mu(x - h_n u)\mu(x - h_n v) \times f_{X,i,j,k}(x, x - h_n u, x - h_n v)\,dx\,dK(u,v) \tag{22}$$
$$= \frac{1}{n^2} \sum_{j \neq k; j,k \neq i} \phi_{nj}\phi_{nk} \int_{-\infty}^{\infty} \mu(x - h_n u)\mu(x - h_n v)\{f_{X,i,j,k}(x, x - h_n u, x - h_n v) - f_X(x) f_{X,j,k}(x - h_n u, x - h_n v)\}\,dx\,dK(u,v)$$
$$+ \frac{1}{n^2} \sum_{j \neq k; j,k \neq i} \phi_{nj}\phi_{nk} \int_{-\infty}^{\infty} \mu(x - h_n u)\mu(x - h_n v) f_X(x) \times \{f_{X,j,k}(x - h_n u, x - h_n v) - f_X(x - h_n u) f_X(x - h_n v)\}\,dx\,dK(u,v)$$

$$
\begin{aligned}
&+\frac{1}{n^2}\sum_{j\neq k;\, j,k\neq i}\phi_{nj}\phi_{nk}\int_{-\infty}^{\infty}\mu(x-h_nu)\mu(x-h_nv)\\
&\qquad\qquad\times f_X(x)f_X(x-h_nu)f_X(x-h_nv)dxdK(u,v)\\
&\le Cn^{-2}\Big[n+\sum_{|j-i|>N}\Big(|j-i|^{2d_X-1}+o(|j-i|^{2d_X-1})\Big)\\
&\qquad+\sum_{|k-i|>N}\Big(|k-i|^{2d_X-1}+o(|k-i|^{2d_X-1})\Big)\\
&\qquad+\sum_{|j-k|>N}\Big(|j-k|^{2d_X-1}+o(|j-k|^{2d_X-1})\Big)\Big]\\
&\qquad+\bar{\phi}^2\int_{-\infty}^{\infty}\mu^2(x)f_X^3(x)dx(1+O(h_n))+O(n^{-1})\\
&=O(n^{2d_X-1}),\quad \text{not depending on } i.
\end{aligned}
$$

This together with (12) and Assumption (B4) implies that

$$
\begin{aligned}
&\sum_{i=1}^{n}\frac{w_i}{f_X^2(X_i)}\Big(\frac{1}{nh_n}\sum_{j=1}^{n}\mu(X_j)\phi_{nj}\tilde{K}_j(X_i)\Big)^2 \qquad (23)\\
&=O_p(\max\{n^{-1}h_n^{-1}\log^2 n,\ n^{2d_X-1}\log^2 n\}).
\end{aligned}
$$

Now consider the first term in the right hand side of (18). Let

$$
D_{ij}=\frac{1}{h_n}\mathrm{E}\Big(\frac{\tilde{K}_j(X_i)}{f_X(X_i)}\Big)-1,\quad 1\le j\le n.
$$

We have

$$
\begin{aligned}
&\frac{1}{nh_n}\sum_{j=1}^{n}\varepsilon_j(1+\phi_{nj})\frac{\tilde{K}_j(X_i)}{f_X(X_i)}-Z_n \qquad (24)\\
&=\frac{1}{n}\frac{\tilde{K}(0)}{h_nf_X(X_i)}\varepsilon_i(1+\phi_{ni})+\frac{1}{nh_n}\sum_{j\neq i}\varepsilon_j(1+\phi_{nj})\frac{\tilde{K}_j(X_i)}{f_X(X_i)}-Z_n\\
&=\frac{1}{n}\Big(\frac{\tilde{K}(0)}{h_nf_X(X_i)}-1\Big)\varepsilon_i(1+\phi_{ni})+\frac{1}{n}\sum_{j\neq i}\varepsilon_j(1+\phi_{nj})D_{ij}\\
&\qquad+\frac{1}{nh_n}\sum_{j\neq i}\varepsilon_j(1+\phi_{nj})\Big[\frac{\tilde{K}_j(X_i)}{f_X(X_i)}-\mathrm{E}\Big(\frac{\tilde{K}_j(X_i)}{f_X(X_i)}\Big)\Big]\\
&=A_{0i}+A_{1i}+A_{2i},\quad \text{say.}
\end{aligned}
$$

First, by (12) and Assumption (B4),

$$\sum_{i=1}^{n} w_i A_{0i}^2 = \frac{1}{n^2}\sum_{i=1}^{n} w_i\Big(\frac{\tilde{K}(0)}{h_n f_X(X_i)} - 1\Big)^2 \varepsilon_i^2(1+\phi_{ni})^2$$
$$= O_p(n^{-2}h_n^{-2}\log^2 n). \tag{25}$$

Next, consider the term A_{1i}. Note that, for $j \neq i$,

$$\begin{aligned}
D_{ij} &= \int_{-\infty}^{\infty}\int_{\tau_{1n}}^{\tau_{2n}} \frac{\tilde{K}(u)}{f_X(x)} f_{X,i,j}(x, x-h_n u)dudx - 1 \\
&= \int_{-\infty}^{\infty}\int_{\tau_{1n}}^{\tau_{2n}} \frac{\tilde{K}(u)}{f_X(x)} f_X(x) f_X(x-h_n u)dudx - 1 \\
&\quad + \int_{-\infty}^{\infty}\int_{\tau_{1n}}^{\tau_{2n}} \frac{\tilde{K}(u)}{f_X(x)}\big[f_{X,i,j}(x, x-h_n u) - f_X(x) f_X(x-h_n u)\big]dudx \\
&= O(h_n) + \int_{-\infty}^{\infty}\int_{\tau_{1n}}^{\tau_{2n}} \frac{\tilde{K}(u)}{f_X(x)}\big[f_{X,i,j}(x, x-h_n u) - f_X(x) f_X(x-h_n u)\big]dudx.
\end{aligned}$$

Then

$$\begin{aligned}
\mathrm{E}(A_{1i}^2) &= \frac{1}{n^2}\sum_{j\neq i}(1+\phi_{nj})^2 D_{ij}^2 \mathrm{E}\varepsilon_j^2 + \frac{1}{n^2}\sum_{j\neq k;\, j,k\neq i}(1+\phi_{nj})(1+\phi_{nk}) D_{ij} D_{ik}\mathrm{E}(\varepsilon_j\varepsilon_k) \\
&= A_{11i} + A_{12i}, \qquad \text{say}.
\end{aligned}$$

But,

$$\begin{aligned}
A_{11i} &\leq \frac{C}{n^2}\sum_{j\neq i} D_{ij}^2 = \frac{C}{n^2}\Big(\sum_{0<|j-i|\leq N} + \sum_{|j-i|>N}\Big) D_{ij}^2 \\
&\leq O(n^{-2}) + \frac{1}{n^2}\sum_{|j-i|>N}\Big(O(h_n^2) + C\Big(|j-i|^{2d_X-1} + o(|j-i|^{2d_X-1})\Big)^2\Big) \\
&= O(n^{-2}) + O(n^{-1}h_n^2) + O(n^{4d_X-3}), \qquad \forall\, 1 \leq i \leq n, \\
\sum_{i=1}^{n} A_{11i} &= O(n^{-1}) + O(h_n^2) + O(n^{4d_X-2}).
\end{aligned}$$

For the sake of brevity, we omit the $o(|j-i|^{2d_X-1})$, $o(|k-i|^{2d_X-1})$ and $o(|j-k|^{2d_X-1})$ terms in the following derivations.

$$A_{12i} \leq \frac{C}{n^2}\sum_{j\neq k;\, j,k\neq i} |D_{ij} D_{ik}||\gamma(j-k)|$$

$$\begin{aligned}
&\le \frac{C}{n^2}\Big[n + \sum_{|j-k|>N;\, j,k\neq i} \Big(h_n + 1 + |i-j|^{2d_X-1}\Big) \\
&\qquad\qquad \times \Big(h_n + |i-k|^{2d_X-1}\Big)|j-k|^{2d-1}\Big] \\
&\le Cn^{-1} + \frac{C}{n^2} \sum_{|j-k|>N;\, j,k\neq i} \Big\{h_n|j-k|^{2d-1} + |j-k|^{2d-1}|i-k|^{2d_X-1} \\
&\quad + h_n|j-k|^{2d-1}|i-j|^{2d_X-1} + |j-k|^{2d-1}|i-j|^{2d_X-1}|i-k|^{2d_X-1}\Big\} \\
&= O(n^{2d-1}h_n) + O(n^{2d+2d_X-2}) + O(n^{2d+4d_X-3}), \qquad \forall\, 1 \le i \le n, \\
\sum_{i=1}^n A_{12i} &= O(n^{2d}h_n) + O(n^{2d+2d_X-1}) + O(n^{2d+4d_X-2}).
\end{aligned}$$

Hence, by (12) and the fact that $n^{2d+2d_X-1}/n^{2d}h_n \to 0$,

$$\sum_{i=1}^n w_i A_{1i}^2 = O_p(n^{2d-1}h_n). \tag{26}$$

Next, we shall show that

$$\sum_{i=1}^n w_i A_{2i}^2 = O_p(\max\{n^{-1}h_n^{-1}\log^2 n, n^{2d-1}h_n\}). \tag{27}$$

Note that

$$\begin{aligned}
A_{2i}^2 &= \frac{1}{n^2h_n^2}\sum_{j\neq i} \varepsilon_j^2(1+\phi_{nj})^2\Big(\frac{\tilde{K}_j(X_i)}{f_X(X_i)} - \mathrm{E}\Big(\frac{\tilde{K}_j(X_i)}{f_X(X_i)}\Big)\Big)^2 \\
&\quad + \frac{1}{n^2h_n^2}\sum_{j\neq k;\, j,k\neq i} (1+\phi_{nj})(1+\phi_{nk})\varepsilon_j\varepsilon_k \\
&\qquad \times \Big(\frac{\tilde{K}_j(X_i)}{f_X(X_i)} - \mathrm{E}\Big(\frac{\tilde{K}_j(X_i)}{f_X(X_i)}\Big)\Big)\Big(\frac{\tilde{K}_k(X_i)}{f_X(X_i)} - \mathrm{E}\Big(\frac{\tilde{K}_k(X_i)}{f_X(X_i)}\Big)\Big) \\
&=: A_{21i} + A_{22i}, \qquad \text{say.}
\end{aligned}$$

Fix an $1 \le i \le n$. By Assumption ($\tilde{\mathrm{K}}$), we obtain

$$\begin{aligned}
&\mathrm{E}\Big(\frac{1}{n^2h_n^2}\sum_{j\neq i}\varepsilon_j^2(1+\phi_{nj})^2\tilde{K}_j^2(X_i)\Big) \\
&= n^{-2}h_n^{-1}\mathrm{E}(\varepsilon_1^2)\sum_{j\neq i}(1+\phi_{nj})^2\int_{-\infty}^{\infty}\tilde{K}^2(u)f_{X,i,j}(x, x-h_nu)dudx = O(n^{-1}h_n^{-1}).
\end{aligned}$$

Then by (12) and Assumption (B4),

$$\sum_{i=1}^{n} w_i A_{21i} = O_p(n^{-1}h_n^{-1}\log^2 n).$$

Moreover, by the independence of $\{\varepsilon_j\}$ and $\{X_j\}$,

$$\big|\mathrm{E}(A_{22i})\big| \le Cn^{-2}h_n^{-2} \sum_{j\neq k;\, j,k\neq i} \Big|\mathrm{E}(\varepsilon_j\varepsilon_k)\mathrm{Cov}\Big(\frac{\tilde{K}_j(X_i)}{f_X(X_i)}, \frac{\tilde{K}_k(X_i)}{f_X(X_i)}\Big)\Big|.$$

Arguing as for (22), by Assumptions (B2) and (B3),

$$\begin{aligned}
&h_n^{-2} \sum_{j\neq k;\, j,k\neq i} \Big|\mathrm{E}(\varepsilon_j\varepsilon_k)\mathrm{Cov}\Big(\frac{\tilde{K}_j(X_i)}{f_X(X_i)}, \frac{\tilde{K}_k(X_i)}{f_X(X_i)}\Big)\Big| \\
&\le \sum_{j\neq k;\, j,k\neq i} |\gamma(j-k)| \Big| \int_{-\infty}^{\infty} \frac{\tilde{K}(u)\tilde{K}(v)}{f_X^2(x)} f_{X,i,j,k}(x, x-h_n u, x-h_n v)dxdudv \\
&\qquad - \int_{-\infty}^{\infty} \frac{\tilde{K}(u)}{f_X(x)} f_{X,i,j}(x, x-h_n u)dxdu \int_{-\infty}^{\infty} \frac{\tilde{K}(v)}{f_X(x)} f_{X,i,k}(x, x-h_n v)dxdv \Big| \\
&\le C\Big[n + \sum_{|j-k|>N;\, j,k\neq i} |j-k|^{2d-1}\Big\{C'|i-j|^{2d_X-1} + C'|i-k|^{2d_X-1} \\
&\qquad + \int_{-\infty}^{\infty} \frac{\tilde{K}(u)\tilde{K}(v)}{f_X^2(x)} f_X(x) f_{X,j,k}(x-h_n u, x-h_n v)dxdudv \\
&\qquad - \Big(C|i-j|^{2d_X-1} + \int_{-\infty}^{\infty} \frac{\tilde{K}(u)}{f_X(x)} f_X(x) f_X(x-h_n u)dxdu\Big) \\
&\qquad\quad \times \Big(C|i-k|^{2d_X-1} + \int_{-\infty}^{\infty} \frac{\tilde{K}(v)}{f_X(x)} f_X(x) f_X(x-h_n v)dxdv\Big)\Big\}\Big] \\
&\le O(n) + C \sum_{|j-k|>N;\, j,k\neq i} |j-k|^{2d-p}\Big\{C'|i-j|^{2d_X-1} + C'|i-k|^{2d_X-1} \\
&\qquad + C|j-k|^{2d_X-1} + \int_{-\infty}^{\infty} \frac{\tilde{K}(u)\tilde{K}(v)}{f_X(x)} f_X(x-h_n u) f_X(x-h_n v)dxdudv \\
&\qquad - \Big(C|i-j|^{2d_X-1} + \int_{-\infty}^{\infty} \tilde{K}(u) f_X(x-h_n u)dxdu\Big) \\
&\qquad\quad \times \Big(C|i-k|^{2d_X-1} + \int_{-\infty}^{\infty} \tilde{K}(v) f_X(x-h_n v)dxdv\Big)\Big\} \\
&\le O(n) + C\Big\{ \sum_{|j-k|>N;\, j,k\neq i} |j-k|^{2d-1}|i-j|^{2d_X-1} + \sum_{|j-k|>N} |j-k|^{2(d+d_X)-2}
\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{|j-k|>N;\, j,k\neq i} |j-k|^{2d-1}|i-k|^{2d_X-1} + \sum_{|j-k|>N} |j-k|^{2d-1} O(h_n) \\
&+ \sum_{|j-k|>N;\, j,k\neq i} |j-k|^{2d-1}|i-j|^{2d_X-1}|i-k|^{2d_X-1}\Big\} \\
= {} & O(n^{2(d+d_X)}) + O(n^{2d+4d_X-1}) + O(n^{2d+1}h_n) = O(n^{2d+1}h_n).
\end{aligned}
$$

This implies the claim in (27).

Assumptions (B5′) combined with (24)–(27) yield

$$
\sum_{i=1}^{n} w_i \Big(\frac{1}{nh_n}\sum_{j=1}^{n} \varepsilon_j (1+\phi_{nj}) \frac{\tilde{K}_j(X_i)}{f_X(X_i)} - Z_n\Big)^2 = O_p(\max\{n^{-1}h_n^{-1}\log^2 n,\ n^{2d-1}h_n\}).
$$

Claim (17) is now established upon combining this bound with (18), (21), and (23).

Now we begin the proof of Lemma 2.

Proof Recall Assumption (B4) and let $I_n := [\tau_{1n},\ \tau_{2n}]$. Clearly,

$$
\begin{aligned}
\max_{1\le i\le n;\, X_i\in I_n} \Big|1 - \frac{f_{n,X}(X_i)}{f_X(X_i)}\Big| &\le \max_{1\le i\le n;\, X_i\in I_n} |f_{n,X}(X_i) - f_X(X_i)| \cdot \max_{1\le i\le n;\, X_i\in I_n} 1/f_X(X_i) \\
&= O_p(\max\{n^{d_X-1/2}h_n^{-1}\log n,\ \ h_n \log n\}).
\end{aligned}
$$

The last claim above follows from Lemma 1 and Assumption (B4), which imply that the second factor above is $O_p(\log(n))$. Hence, by (12),

$$
\max_{1\le i\le n} \Big|1 - \frac{f_{n,X}(X_i)}{f_X(X_i)}\Big|^2 w_i = O_p(\max\{n^{2d_X-2}h_n^{-2}\log^2 n,\ \ n^{-1}h_n^2 \log^2 n\}).
$$

Recall the results from (8) and (9),

$$
Z_n = \bar{\varepsilon}_n + \bar{W}_n = O_p(n^{d-1/2}). \tag{28}
$$

Thus it follows from Lemma 3 that $\sum_{i=1}^n (\hat{Z}_i - Z_n)^2 w_i$ is bounded from the above by 2 times the sum

$$
\begin{aligned}
&\sum_{i=1}^{n} \Big(\frac{f_X(X_i)}{f_{n,X}(X_i)}\Big)^2 \Big(\frac{f_{n,X}(X_i)}{f_X(X_i)}\hat{Z}_i - Z_n\Big)^2 w_i + \sum_{i=1}^{n} \Big(\frac{f_X(X_i)}{f_{n,X}(X_i)}\Big)^2 \Big(1 - \frac{f_{n,X}(X_i)}{f_X(X_i)}\Big)^2 w_i Z_n^2 \\
&= O_p(\gamma_n^2) + O_p(n^{2d}) O_p(\max\{n^{2d_X-2}h_n^{-2}\log^2 n,\ \ n^{-1}h_n^2\log^2 n\}) = O_p(\xi_n^2).
\end{aligned}
$$

This completes the proof of Lemma 2.

We also need the following uniform reduction lemma.

Lemma 4 *With w_i's as specified in (5) and under assumptions (A) and (B3),*

$$\sup_{x\in\mathbb{R}} \Big| \sum_{i=1}^{n} w_i [I(\varepsilon_i \le x) - F(x) + f(x)\varepsilon_i] \Big| = o_p(n^{d-1/2}).$$

The proof of this lemma is very similar to that of Theorem 10.2.3 of Giraitis et al. [5], which uses a chaining argument that heavily depends on Lemma 10.2.5 dealing only with terms involving ε_i's. One uses the same chaining argument and Lemma 10.2.5 together with the independence of $\{X_i\}$ and $\{\varepsilon_j\}$ and the fact that $0 \le w_i \le 1,\ 1 \le i \le n$, and $\sum_{i=1}^{n} w_i = 1$, to prove the above result. We leave out the details for an interested reader. The only difference between this Lemma and Theorem 10.2.3 is that there the weights γ_{nj} are nonrandom whereas here w_i are random, but independent of ε_i's.

Now we begin the proof of Theorem 1.

Proof of Theorem 1 Let

$$G(x) = \int_{-\infty}^{x} K(u)du, \quad F_n(x) = n^{-1}\sum_{i=1}^{n} I(\varepsilon_i \le x), \quad F_{nw}(x) = \sum_{i=1}^{n} w_i I(\varepsilon_i \le x).$$

Recall $\hat{Z}_i = \varepsilon_i - \hat{\varepsilon}_i$ and $Z_n = \bar{\varepsilon}_n + \bar{W}_n$ from (16), and let $\varDelta_i = \hat{Z}_i - Z_n$. By the definition of $\hat{F}(x)$, we have

$$\begin{aligned}
&\hat{F}(x) - F(x) - \bar{W}_n f(x)\\
&= \sum_{i=1}^{n} w_i \Big[G\Big(\frac{x-\hat{\varepsilon}_i}{b_n}\Big) - F(x) - \bar{W}_n f(x)\Big]\\
&= \sum_{i=1}^{n} w_i \Big[G\Big(\frac{x+\hat{Z}_i-\varepsilon_i}{b_n}\Big) - G\Big(\frac{x+Z_n-\varepsilon_i}{b_n}\Big)\Big]\\
&\qquad + \int_{-\infty}^{\infty} \Big[G\Big(\frac{x+Z_n-u}{b_n}\Big) - F(x) - \bar{W}_n f(x)\Big] dF_{nw}(u)\\
&= \sum_{i=1}^{n} w_i \Big[K\Big(\frac{x+Z_n-\varepsilon_i}{b_n}\Big)\frac{\varDelta_i}{b_n} + R_i(x)\Big]\\
&\qquad + \int_{-\infty}^{\infty} \Big[G\Big(\frac{x+Z_n-u}{b_n}\Big) - F(x) - \bar{W}_n f(x)\Big] dF_{nw}(u)\\
&= \sum_{i=1}^{n} w_j K\Big(\frac{x+Z_n-\varepsilon_i}{b_n}\Big)\frac{\varDelta_i}{b_n} + \sum_{i=1}^{n} w_i R_i(x)\\
&\qquad + \int_{-\infty}^{\infty} \Big[G\Big(\frac{x+Z_n-u}{b_n}\Big) - F(x) - \bar{W}_n f(x)\Big] dF_{nw}(u)\\
&=: I_1(x) + I_2(x) + I_3(x), \qquad \text{say.}
\end{aligned}$$

Here

$$R_i(x) := G\Big(\frac{x+\hat{Z}_i-\varepsilon_i}{b_n}\Big) - G\Big(\frac{x+Z_n-\varepsilon_i}{b_n}\Big) - \frac{\Delta_i}{b_n}G'\Big(\frac{x+Z_n-\varepsilon_i}{b_n}\Big)$$
$$= \int_0^{\Delta_i/b_n} \Big(\frac{\Delta_i}{b_n}-s\Big)K'\Big(\frac{x+Z_n-\varepsilon_i}{b}+s\Big)ds.$$

Note that

$$|w_iR_i(x)| \le \frac{|w_i\Delta_i|}{b_n}\int_{-|\Delta_i|/b_n}^{|\Delta_i|/b_n}\Big|K'(\frac{x+Z_n-\varepsilon_i}{b_n}+s)\Big|ds, \quad \forall\, 1\le i\le n, x\in\mathbb{R}.$$

Hence, by Lemma 2,

$$\sup_{x\in\mathbb{R}}|I_2(x)| \le Cb_n^{-2}\sup_{x\in\mathbb{R}}|K'(x)|\sum_{i=1}^n w_i\Delta_i^2 = O_p(\xi_n^2b_n^{-2}) = o_p(n^{d-1/2}).$$

Next, similarly,

$$\sup_{x\in\mathbb{R}}|I_1(x)| \le Cb_n^{-1}\sup_{x\in\mathbb{R}}|K(x)|\sum_{i=1}^n |w_i\Delta_i| = O_p(\xi_nb_n^{-1}) = o_p(n^{d-1/2}).$$

Note that

$$I_3(x) = \int_{-\infty}^{\infty}\big(F_{nw}(x+Z_n-ub_n)-F(x)-\bar{W}_nf(x)\big)K(u)du$$
$$= \int_{-\infty}^{\infty}\big(F_{nw}(x+Z_n-ub_n)-F(x-ub_n)-\bar{W}_nf(x)\big)K(u)du$$
$$+ \int_{-\infty}^{\infty}\big(F(x-ub_n)-F(x)\big)K(u)du$$
$$= \sum_{i=1}^3\psi_{ni}(x)+O(b_n^2),$$

where

$$\psi_{n1}(x) = \int_{-\infty}^{\infty}\big(F_{nw}(x+Z_n-ub_n)-F(x+Z_n-ub_n)+f(x+Z_n-ub_n)\bar{\varepsilon}_n\big)K(u)du,$$
$$\psi_{n2}(x) = \int_{-\infty}^{\infty}\big(F(x+Z_n-ub_n)-F(x-ub_n)-f(x+Z_n-ub_n)Z_n\big)K(u)du,$$
$$\psi_{n3}(x) = \bar{W}_n\int_{-\infty}^{\infty}(f(x+Z_n-ub_n)-f(x))K(u)du.$$

Note that for any $y \in \mathbb{R}$,

$$F_{nw}(y) - F(y) + f(y)\bar{\varepsilon}_n = \sum_{i=1}^{n} w_i[I(\varepsilon_i \le y) - F(y) + f(y)\varepsilon_i] + f(y)\sum_{i=1}^{n}(n^{-1} - w_i)\varepsilon_i.$$

From (13), we obtain

$$\begin{aligned}
&\mathrm{E}\Big(\sum_{i=1}^{n} \bar{w}(n^{-1} - w_i)\varepsilon_i\Big)^2 \\
&= \sum_{i=1}^{n} \mathrm{E}|\bar{w}(n^{-1} - w_i)|^2 \mathrm{E}\varepsilon_i^2 + \sum_{i \ne j} \mathrm{E}\big\{\bar{w}^2(n^{-1} - w_i)(n^{-1} - w_j)\big\}\gamma(i-j) \\
&\le o(n^{-1}) + \max_{1 \le i \le n} \mathrm{E}\big|\bar{w}(n^{-1} - w_i)\big|^2 \sum_{i \ne j} |\gamma(i-j)| \\
&= o(n^{-1}) + o(n^{-2})O(n^{2d+1}) = o(n^{2d-1}).
\end{aligned}$$

This, together with (12), implies that

$$\sum_{i=1}^{n}(n^{-1} - w_i)\varepsilon_i = o_p(n^{d-1/2}).$$

Hence by Lemma 4,

$$\sup_{x \in \mathbb{R}} |\psi_{n1}(x)| = o_p(n^{d-1/2}).$$

Next consider the ψ_{n2} term. Following the proof of Theorem 2.2 of Koul et al. [7], we obtain

$$\psi_{n2}(x) = -\int_0^{Z_n}\int_v^{Z_n}\int_{-\infty}^{\infty} f'(x - ub_n + w)K(u)\,du\,dw\,dv.$$

Therefore,

$$\sup_{x \in \mathbb{R}} |\psi_{n2}(x)| \le CZ_n^2 = O_p(n^{2d-1}) = o_p(n^{d-1/2}).$$

Finally,

$$\sup_{x \in \mathbb{R}} |\psi_{n3}(x)| \le |\bar{W}_n||CZ_n + Cb_n| = O_p(n^{2d-1}) + O(b_n n^{d-1/2}) = o_p(n^{d-1/2}).$$

This concludes the proof of (10), and hence of Theorem 1.

3 Estimation of d and $c(\phi)$

In practice, in order to be able to use the studentized version of the estimator $\hat{F}$ to make the large sample inference about F, or to use the uniform reduction principle to propose a goodness of fit test for F, a $\log n$-consistent estimator of d and a consistent estimator of $c(\phi)$ are needed. Several approaches for estimating the long memory parameter d have been suggested in the literature. For a recent review on the estimators of the long memory parameter, see Giraitis et al. [5]. For semiparametric models, popular estimators are the local Whittle estimators studied by Robinson [16–18]. Proceeding as in Robinson [18], let $\lambda_j = 2\pi j/n$, and define

$$Q_j = \frac{1}{2\pi n}\Big|\sum_{t=1}^{n} \omega_t \hat{\varepsilon}_t \exp(\mathbf{i}\, t\lambda_j)\Big|^2, \qquad \tilde{Q}_j = \frac{1}{2\pi n}\Big|\sum_{t=1}^{n} \varepsilon_t \exp(\mathbf{i}\, t\lambda_j)\Big|^2,$$

where $\omega_t = w(X_t)$ with w as in (5). Fix $0 < r_1 < r_2 < 1/2$. With an integer $m \in [1, n/2)$, for $r_1 \le \psi \le r_2$, let

$$R(\psi) = \log\Big(\frac{1}{m}\sum_{j=1}^{m} \lambda_j^{2\psi-1} Q_j\Big) - (2\psi - 1)\sum_{j=1}^{m} \log\lambda_j.$$

As pointed out in Robinson [17], the spectral density of the error process $\{\varepsilon_i,\ i \in \mathbb{Z}\}$ satisfies $g(\lambda) \sim G\lambda^{-2d}$, as $\lambda \to 0_+$, with G being some positive constant. Let $r_1 \le d \le r_2$. Then the local Whittle estimator of d based on the residuals $\{\hat{\varepsilon}_i\}$ is defined as

$$\hat{d} = \mathrm{argmin}_{\psi\in[r_1,r_2]} R(\psi).$$

Theorem 2 *In addition to the assumptions of Lemma 2, we assume that* $m \to \infty$, $nm^{-2}\log m = o(1)$, $m^{2d} n^{2d_X - 1}\log m = o(1)$, *and in a neighborhood of the origin,* $a(\lambda) = \sum_{j=0}^{\infty} a_j e^{\mathbf{i}\, j\lambda}$ *is differentiable and* $da(\lambda)/d\lambda = O(|a(\lambda)|/\lambda)$ *as* $\lambda \to 0_+$. *Then,* $(\log n)(\hat{d} - d) \to_p 0$.

To prove Theorem 2, we need the following lemma. Let $D_j = (Q_j - \tilde{Q}_j)/g(\lambda_j)$, $1 \le j \le m$.

Lemma 5 *Under the conditions of Theorem 2, if the following three claims hold, then* $(\log n)(\hat{d} - d) \to_p 0$.

$$\sum_{i=1}^{m-1}\Big(\frac{i}{m}\Big)^{2(r_1-d)+1}\frac{1}{i^2}\Big|\sum_{j=1}^{i} D_j\Big| = o_p(1), \tag{29}$$

$$(\log n)^2\sum_{i=1}^{m-1}\Big(\frac{i}{m}\Big)^{1-2\delta}\frac{1}{i^2}\Big|\sum_{j=1}^{i} D_j\Big| = o_p(1), \qquad \frac{(\log n)^2}{m}\sum_{j=1}^{m} D_j = o_p(1),$$

where δ is a small positive constant.

Proof The proof is the same as the part (ii) proof of Theorem 3 in Robinson [18].

Proof of Theorem 2 According to Lemma 5, to prove Theorem 2, it suffices to verify the three claims in (29). Since the proofs of the last two claims are similar to the first one, we shall only prove the first claim in (29). Note that

$$|Q_j - \tilde{Q}_j| \le \tilde{Q}_{j,w} + Q_{j,z} + 2|\tilde{Q}_j Q_{j,z}|^{1/2} + 2|\tilde{Q}_j \tilde{Q}_{j,w}|^{1/2} + 2|\tilde{Q}_{j,w} Q_{j,z}|^{1/2}, \qquad (30)$$

where

$$\tilde{Q}_{j,w} = \frac{1}{2\pi n}\Big|\sum_{t=1}^{n}(\omega_t - 1)\varepsilon_t \exp(\mathbf{i}\, t\lambda_j)\Big|^2, \quad Q_{j,z} = \frac{1}{2\pi n}\Big|\sum_{t=1}^{n}\omega_t \hat{Z}_t \exp(\mathbf{i}\, t\lambda_j)\Big|^2.$$

Dalla, Giraitis and Hidalgo [2] proved that

$$\sum_{j=1}^{m}\frac{\tilde{Q}_j}{g(\lambda_j)} = O_p(m). \qquad (31)$$

In addition, from Zygmund ([21], page 90),

$$\Big|\sum_{t=1}^{n} e^{\mathbf{i}\, t\lambda}\Big| \le C|\lambda|^{-1}, \quad 0 < |\lambda| \le \pi. \qquad (32)$$

Lemma 2 with (28) and Assumption (B5′) implies the bound

$$\sum_{t=1}^{n} w_t \hat{Z}_t^2 = O_p(\xi_n^2) + O_p(n^{2d-1}) = O_p(n^{2d-1}). \qquad (33)$$

Since $\omega_t = n\bar{w}\, w_t$, where $\bar{w} = \frac{1}{n}\sum_{t=1}^{n} w(X_t)$, and by (12),

$$\sum_{t=1}^{n}\omega_t^2 \hat{Z}_t^2 \le (n\bar{w})^2 \max_{1\le t\le n}|w_t| \sum_{t=1}^{n} w_t \hat{Z}_t^2 = O_p(n^{2d}).$$

This yields

$$\begin{aligned}\frac{|Q_{j,z}|}{g(\lambda_j)} &\le \frac{Cn^{-1}\Big|\sum_{t=1}^{n}\omega_t^2\hat{Z}_t^2 \sum_{t=1}^{n} e^{\mathbf{i}\, t2\lambda_j}\Big|}{g(\lambda_j)} \\ &\le O_p(n^{2d-1})C\lambda_j^{2d-1} = O_p(j^{2d-1}).\end{aligned} \qquad (34)$$

Next, we consider $\tilde{Q}_{j,w}$. Recall $g(u)$ is the spectral density of the error process $\{\varepsilon_i\}$, $\mathrm{E}(\varepsilon_t \varepsilon_s) = \int_{-\pi}^{\pi} g(u) \exp(-\mathbf{i}\,(t-s)u)du$. Let $J(\lambda) = \sum_{t=1}^{n} \mathrm{E}(\omega_t - 1) \exp(\mathbf{i}\, t\lambda)$. Then

$$
\begin{aligned}
&2\pi n \mathrm{E}(\tilde{Q}_{j,w}) \qquad (35)\\
&= \sum_{t,s=1}^{n} \mathrm{E}\big((\omega_t - 1)(\omega_s - 1)\big) \exp(\mathbf{i}\,(t-s)\lambda_j) \mathrm{E}(\varepsilon_t \varepsilon_s)\\
&= \sum_{t,s=1}^{n} \big(\mathrm{E}(\omega_t - 1)\mathrm{E}(\omega_s - 1) + \mathrm{Cov}(\omega_t, \omega_s)\big) \exp(\mathbf{i}\,(t-s)\lambda_j) \mathrm{E}(\varepsilon_t \varepsilon_s)\\
&= \int_{-\pi}^{\pi} g(u) |J(\lambda_j - u)|^2 du + O(n) + O(n^{2d+2d_X})\\
&\le g(\lambda_j) \int_{-\pi}^{\pi} |J(\lambda_j - u)|^2 du + \Big(\int_{3\lambda_j/2}^{\pi} + \int_{-\pi}^{\lambda_j/2}\Big) g(u) |J(\lambda_j - u)|^2 du\\
&\quad + \int_{\lambda_j/2}^{3\lambda_j/2} |g(u) - g(\lambda_j)| |J(\lambda_j - u)|^2 du + O(n) + O(n^{2d+2d_X})\\
&=: R_1 + R_2 + R_3 + O(n) + O(n^{2d+2d_X}), \quad \text{say.}
\end{aligned}
$$

The claim in the third equality above follows from (14).

By (11) and (32), for $0 < |\lambda| \le \pi$,

$$
|J(\lambda)| \le C \mathrm{E}|\omega_t - 1|/|\lambda| = o(|\lambda|^{-1}).
$$

Then, for some $u \in [-\pi, \pi]$,

$$
R_1 = 2\pi g(\lambda_j) |J(\lambda_j - u)|^2 \le C g(\lambda_j) |\lambda_j - u|^{-2} = O(g(\lambda_j) n^2 j^{-2}).
$$

We split the first part of R_2 into two components, the second part can be treated similarly. For all $d \in (0, 1/2)$ and sufficiently small λ_j, there exists $\varepsilon \in (3\lambda_j/2, \pi)$ such that $g(u)/u^{1/2-d} = O(g(\lambda_j)/\lambda_j^{1/2-d})$, for $\lambda_j < u < \varepsilon$, and

$$
\begin{aligned}
&\int_{3\lambda_j/2}^{\pi} g(u) |J(\lambda_j - u)|^2 du\\
&= \int_{3\lambda_j/2}^{\varepsilon} g(u) |J(\lambda_j - u)|^2 du + \int_{\varepsilon}^{\pi} g(u) |J(\lambda_j - u)|^2 du\\
&= O\Big(\frac{g(\lambda_j)}{\lambda_j^{1/2-d}} \int_{3\lambda_j/2}^{\varepsilon} \frac{u^{1/2-d}}{(\lambda_j - u)^2} du + \frac{1}{(\lambda_j - \varepsilon)^2} \int_{-\pi}^{\pi} g(u) du\Big)\\
&= O(g(\lambda_j)\lambda_j^{-1} + \varepsilon^{-2}) = O(g(\lambda_j)\lambda_j^{-1}).
\end{aligned}
$$

Hence

$$R_2 = O(g(\lambda_j)nj^{-1}).$$

Note that $g(\lambda) = \frac{1}{2\pi}\big|\sum_{j=0}^{\infty} a_j e^{\mathbf{i}j\lambda}\big|^2$. Then by the assumption,

$$\sup_{\lambda_j/2\le u\le 3\lambda_j/2} |g(\lambda_j) - g(u)|/|\lambda_j - u| = O(g(\lambda_j)/\lambda_j),$$

so that R_3 is bounded by

$$\begin{aligned}
&\frac{Cg(\lambda_j)}{\lambda_j}\int_{\lambda_j/2}^{3\lambda_j/2} |J(\lambda_j - u)|du \le \frac{Cg(\lambda_j)}{\lambda_j}\int_0^{\lambda_j/2} |J(\mu)|du \\
&= \frac{Cg(\lambda_j)}{\lambda_j}\Big(\int_0^{1/n} + \int_{1/n}^{\lambda_j/2}\Big)|J(\mu)|du \le \frac{Cg(\lambda_j)}{\lambda_j} O\Big(\frac{1}{n}n + \log j\Big) \\
&= O(g(\lambda_j)nj^{-1}(1+\log j)).
\end{aligned}$$

These bounds, together with (35), imply that

$$\frac{\tilde{Q}_{j,w}}{g(\lambda_j)} = O_p(j^{-1}\log j + nj^{-2} + j^{2d}n^{-2d} + j^{2d}n^{2d_X-1}). \tag{36}$$

By changing the order of summation, the left hand side of (29) is bounded above by $Cm^{-2(r_1-d)-1}\sum_{j=1}^m j^{2(r_1-d)}|D_j|$, for $d > r_1$, and by $Cm^{-1}\log m\sum_{j=1}^m |D_j|$, for $d = r_1$. Combining (30), (31), (34) and (36) yields, for $d > r_1$,

$$\begin{aligned}
&Cm^{-2(r_1-d)-1}\sum_{j=1}^m j^{2(r_1-d)}|D_j| \\
&\le Cm^{-1}\Bigg\{\sum_{j=1}^m \frac{\tilde{Q}_{j,w}}{g(\lambda_j)} + \sum_{j=1}^m \frac{Q_{j,z}}{g(\lambda_j)} + \Bigg(\sum_{j=1}^m \frac{\tilde{Q}_j}{g(\lambda_j)}\sum_{j=1}^m \frac{Q_{j,z}}{g(\lambda_j)}\Bigg)^{1/2} \\
&\qquad + \Bigg(\sum_{j=1}^m \frac{\tilde{Q}_j}{g(\lambda_j)}\sum_{j=1}^m \frac{\tilde{Q}_{j,w}}{g(\lambda_j)}\Bigg)^{1/2} + \Bigg(\sum_{j=1}^m \frac{\tilde{Q}_{j,w}}{g(\lambda_j)}\sum_{j=1}^m \frac{Q_{j,z}}{g(\lambda_j)}\Bigg)^{1/2}\Bigg\} \\
&= O_p(m^{-1}\log m) + o_p(nm^{-2}) + O_p(m^{2d-1}) + O_p(m^{2d}n^{-2d}) + O_p(m^{2d}n^{2d_X-1}) \\
&= o_p(1).
\end{aligned}$$

Similarly, for $d = r_1$,

$$Cm^{-1}\log m\sum_{j=1}^m |D_j| = o_p(1).$$

This completes the proof of Theorem 2.

Now we turn to estimate $c(\phi)$. Note that

$$c^2(\phi) = \lim_{n\to\infty} n^{-2d-1} \sum_{j,k=1}^{n} \phi_{nj}\phi_{nk}\gamma(j-k).$$

This suggests the estimator

$$\hat{c}^2(\phi) = q^{-2\hat{d}-1} \sum_{j,k=1}^{q} \phi_{qj}\phi_{qk}\hat{\gamma}(j-k),$$

based on the sample auto-covariance function of the residuals, where $q \to \infty$, $q = o(n)$ is a bandwidth sequence, and $\hat{\gamma}(k) = n\sum_{i=1}^{n} w_i w_{i+k}\hat{\varepsilon}_i\hat{\varepsilon}_{i+k}$ (here we again use the random weights w_i defined in (5) to estimate $\gamma(k)$). Theorem 3 below shows that $\hat{c}^2(\phi)$ is a consistent estimator of $c^2(\phi)$.

Theorem 3 *In addition to the assumptions imposed in Lemma 2, we assume that* $E|\zeta_0|^3 < \infty$ *and* $q = o(n^{1/2})$. *Then* $\hat{c}^2(\phi) \to_p c^2(\phi)$.

Proof As discussed above, $\hat{d}$ is $\log(n)$-consistent estimator of d. Hence, to prove Theorem 3, it suffices to show that

$$\tilde{c}^2(\phi) \to_p c^2(\phi),$$

where $\tilde{c}^2(\phi) = q^{-2d-1}\sum_{j,k=1}^{q} \phi_{qj}\phi_{qk}\hat{\gamma}(j-k)$.

We will follow the argument as in the proof of Lemma 2.2 of Koul and Surgailis [10]. Write $\tilde{c}^2(\phi) = \tilde{c}_1^2(\phi) + \tilde{c}_2^2(\phi)$, $\tilde{c}_i^2(\phi) = q^{-2d-1}\sum_{j,k=1}^{q} \phi_{qj}\phi_{qk}\hat{\gamma}_i(j-k)$, $i = 1, 2$, where

$$\hat{\gamma}_1(k) = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i\varepsilon_{i+k}, \qquad \hat{\gamma}_2(k) = \hat{\gamma}(k) - \hat{\gamma}_1(k).$$

By Lemma 2.2 of Koul and Surgailis [10], $\tilde{c}_1^2(\phi) \to_p c^2(\phi)$. Hence, Theorem 3 will follow if we show that

$$nq^{-2d-1} \sum_{j,k=1}^{q}\sum_{i=1}^{n} \phi_{qj}\phi_{qk}\varepsilon_i\varepsilon_{i+j-k}(w_i w_{i+j-k} - n^{-2}) = o_p(1), \tag{37}$$

$$nq^{-2d-1} \sum_{j,k=1}^{q}\sum_{i=1}^{n} \phi_{qj}\phi_{qk}\varepsilon_i w_i \hat{Z}_{i+j-k} w_{i+j-k} = o_p(1), \tag{38}$$

$$nq^{-2d-1} \sum_{j,k=1}^{q}\sum_{i=1}^{n} \phi_{qj}\phi_{qk}\varepsilon_{i+j-k} w_{i+j-k}\hat{Z}_i w_i = o_p(1), \tag{39}$$

$$nq^{-2d-1}\sum_{j,k=1}^{q}\sum_{i=1}^{n}\phi_{qj}\phi_{qk}\hat{Z}_i w_i \hat{Z}_{i+j-k}w_{i+j-k} = o_p(1). \tag{40}$$

Write the left hand side of (37) as

$$nq^{-2d-1}\sum_{j,k=1}^{q}\sum_{i=1}^{n}\phi_{qj}\phi_{qk}\gamma(j-k)(w_i w_{i+j-k} - n^{-2}) + R,$$

where

$$R = nq^{-2d-1}\sum_{j,k=1}^{q}\sum_{i=1}^{n}\phi_{qj}\phi_{qk}\big(\varepsilon_i\varepsilon_{i+j-k} - \gamma(j-k)\big)(w_i w_{i+j-k} - n^{-2}).$$

The facts (12) and (13) imply that

$$\begin{aligned}\max_{1\le i,j,k\le n}|w_i w_{i+j-k} - n^{-2}| &\le \max_{1\le i,j,k\le n}|w_{i+j-k}(w_i - n^{-1}) + n^{-1}(w_{i+j-k} - n^{-1})| \\ &= o_p(n^{-2}).\end{aligned}$$

Then it follows from the proof of Lemma 2.2 of Koul and Surgailis [10] that

$$R = o_p(1).$$

These facts yield

$$\begin{aligned}&nq^{-2d-1}\sum_{j,k=1}^{q}\sum_{i=1}^{n}\phi_{qj}\phi_{qk}\varepsilon_i\varepsilon_{i+j-k}(w_i w_{i+j-k} - n^{-2}) \\ &\le Cn^2q^{-2d-1}\max_{1\le i,j,k\le n}|w_i w_{i+j-k} - n^{-2}|\sum_{j,k=1}^{q}|\gamma(j-k)| + o_p(1) \\ &\le\ O_p(n^2q^{-2d-1})o_p(n^{-2})O(q^{2d+1}) + o_p(1) = o_p(1).\end{aligned}$$

Next, (33) implies

$$\begin{aligned}&nq^{-2d-1}\sum_{j,k=1}^{q}\sum_{i=1}^{n}\phi_{qj}\phi_{qk}\varepsilon_i w_i \hat{Z}_{i+j-k}w_{i+j-k} \\ &\le Cnq^{1-2d}\Big\{\sum_{i=1}^{n}\varepsilon_i^2 w_i^2\sum_{i=1}^{n}\hat{Z}_i^2 w_i^2\Big\}^{1/2} \le Cq^{1-2d}n^{d-1/2} = o_p(1).\end{aligned}$$

This implies (38), while (39) follows in a similar way. Finally, the left hand side of (40) is bounded from the above by

$$Cnq^{1-2d}\sum_{i=1}^{n}\hat{Z}_i^2 w_i^2 \le Cq^{1-2d}n^{2d-1} = o_p(1).$$

This completes the proof of Theorem 3.

Acknowledgements Research of Hira L. Koul was in part supported by the NSF-DMS grant 1205271. Research of Lihong Wang was in part supported by NSFC Grants 11671194 and 11171147. Authors are grateful to a delegant referee whose comments helped to improve the presentation and some of the proofs.

Appendix

In this Appendix, we shall prove that Assumption (B3) implies Assumption (B3′). Let $\phi(u) = \mathrm{E}e^{\mathbf{i}u\eta_0}$, $\hat{f}_X(u) = \mathrm{E}e^{\mathbf{i}uX_0} = \prod_{j\ge 0}\phi(ub_j)$ and let $f_{X,0,t}(x_1, x_2)$ be the joint probability density of (X_0, X_t).

Lemma A. *Under Assumption (B3), the joint probability density* $f_{X,0,t_1,t_2}(x_1, x_2, x_3)$ *of* (X_0, X_{t_1}, X_{t_2}) *exists and*

$$\begin{aligned} &f_{X,0,t_1,t_2}(x_1, x_2, x_3) - f_X(x_1)f_{X,0,\Delta_t}(x_2, x_3) \\ &\quad - \gamma_X(t_1)f'_X(x_1)f'_{X,\Delta_t,x_2}(x_2, x_3) - \gamma_X(t_2)f'_X(x_1)f'_{X,\Delta_t,x_3}(x_2, x_3) = O(t^{2d_X-1-\alpha}), \end{aligned} \tag{41}$$

uniformly in x_1, x_2, x_3 *as* $t_1 \to \infty$, $t_2 \to \infty$ *and* $\Delta_t \to \infty$, *where* $0 < \alpha < \min(d_X/7, (1-2d_X)/8)$, $t = \min(t_1, t_2)$, $\Delta_t = |t_2 - t_1|$, $f'_{X,\Delta_t,x_2}(x_2, x_3) = \partial f_{X,0,\Delta_t}(x_2, x_3)/\partial x_2$ *and* $f'_{X,\Delta_t,x_3}(x_2, x_3) = \partial f_{X,0,\Delta_t}(x_2, x_3)/\partial x_3$.

Proof We extend the proof of Lemma 2 in Giraitis et al. [4] to the trivariate case. We shall first show that for any integer $k \ge 0$,

$$\int_{\mathbb{R}^3} |u|^k|\hat{f}_{X,t_1,t_2}(u)|du = O(1), \quad t_1, t_2, \Delta_t \to \infty, \tag{42}$$

where

$$\begin{aligned}\hat{f}_{X,t_1,t_2}(u) = \hat{f}_{X,t_1,t_2}(u_1, u_2, u_3) &= Ee^{\mathbf{i}(u_1X_0+u_2X_{t_1}+u_3X_{t_2})} \\ &= \prod_j \phi(u_1b_{-j} + u_2b_{t_1-j} + u_3b_{t_2-j})\end{aligned}$$

is the trivariate characteristic function. In particular, (42) implies that the trivariate density $f_{X,0,t_1,t_2}$ exists and belongs to $C^k(\mathbb{R}^3)$, provided t_1, t_2 and Δ_t are large enough. Given $0 < \delta_X < 1$ and $k \ge 0$, there exist disjoint finite sets $J_1, J_2 \in Z$, $|J_1| = |J_2|$, $c_1, c_2 > 0$ independent of t_1, t_2, and $t_0 = t_0(k, \delta_X) > 0$ such that, for any $t_1, t_2, \Delta_t \ge t_0$, $|b_{-j}| > 2|b_{t_1-j}| + c_1$ $(j \in J_1)$, $|b_{-j}| > 2|b_{t_2-j}| + c_2$ $(j \in J_1)$,

$|b_{t_1-j}| > 2|b_{-j}| + c_1$ $(j \in J_2)$, $|b_{t_1-j}| > 2|b_{t_2-j}| + c_2$ $(j \in J_2)$ and $[\delta_X |J_i|] = k + 3$, $i = 1, 2$. By Assumption (B3),

$$\hat{f}_{X,t_1,t_2}(u_1, u_2, u_3) \le C \prod_{J_1 \cup J_2} (1 + |u_1 b_{-j} + u_2 b_{t_1-j} + u_3 b_{t_2-j}|)^{-\delta_X}$$
$$\le C \prod_{J_1} (1 + |u_1 + u_2 r_{1j} + u_3 r_{2j}|)^{-\delta_X} \prod_{J_2} (1 + |u_2 + u_1 r'_{1j} + u_3 r'_{2j}|)^{-\delta_X},$$

where $|r_{ij}| \le 1/2$ $(j \in J_1)$, $|r'_{ij}| \le 1/2$ $(j \in J_2)$, $i = 1, 2$. Note that for any $u_1, u_2, u_3 \in \mathbb{R}$, and any $|r_i| \le 1/2$, $|r'_i| \le 1/2$, $i = 1, 2$,

$$(1 + |u_1 + u_2 r_1 + u_3 r_2|)(1 + |u_2 + u_1 r'_1 + u_3 r'_2|) \ge C(1 + |u|) \equiv C(1 + (u_1^2 + u_2^2 + u_3^2)^{1/2}),$$

where the constant C is independent of r_1, r_2, r'_1, r'_2. Hence we obtain

$$\hat{f}_{X,t_1,t_2}(u_1, u_2, u_3) \le C(1 + |u|)^{-k-3},$$

where the constant $C = C(k)$ does not depend on t_1, t_2. This in turn implies (42).

Write $p_{t_1,t_2}(x_1, x_2, x_3)$ for the left-hand side of (41). Let

$$\hat{p}_{t_1,t_2}(u_1, u_2, u_3) \tag{43}$$
$$= \hat{f}_{X,t_1,t_2}(u_1, u_2, u_3) - \hat{f}_X(u_1) \hat{f}_{X,\Delta_t}(u_2, u_3) + \gamma_X(t_1) u_1 u_2 \hat{f}_X(u_1) \hat{f}_{X,\Delta_t}(u_2, u_3)$$
$$+ \gamma_X(t_2) u_1 u_3 \hat{f}_X(u_1) \hat{f}_{X,\Delta_t}(u_2, u_3),$$

and $\hat{f}_{X,\Delta_t}(u_2, u_3) = Ee^{\mathbf{i}(u_2 X_{t_1} + u_3 X_{t_2})} = \prod_j \phi(u_2 b_{t_1-j} + u_3 b_{t_2-j})$. Then

$$p_{t_1,t_2}(x_1, x_2, x_3) = (2\pi)^{-3} \int_{\mathbb{R}^3} e^{-\mathbf{i}\, x \cdot u} \hat{p}_{t_1,t_2}(u_1, u_2, u_3) du.$$

For any $\alpha > 0$, with $k \ge (1 - 2d_X + \alpha)/\alpha$, we obtain that

$$\int_{\mathbb{R}^3} |\hat{f}_{X,t_1,t_2}(u_1, u_2, u_3)| I(|u| > t^\alpha) du \tag{44}$$
$$\le t^{-k\alpha} \int_{\mathbb{R}^3} |u|^k |\hat{f}_{X,t_1,t_2}(u_1, u_2, u_3)| du \le C t^{-k\alpha} = O(t^{2d_X - 1 - \alpha}).$$

From (2.7) in the proof of Lemma 1 of Giraitis et al. [4], for any $k \ge 0$ and $r > k + 1$,

$$\int_{\mathbb{R}} |u|^k |\hat{f}_X(u)| du \le C \int_{\mathbb{R}} |u|^k / (1 + |u|)^r du = O(1),$$

and (2.13) in the proof of Lemma 2 of Giraitis et al. [4] yields that,

$$\int_{\mathbb{R}^2} |u|^k |\hat{f}_{X,\Delta_t}(u)| du = O(1).$$

Then a similar estimate as in (44) is valid for the three other terms on the right-hand side of (43). Thus

$$\int_{\mathbb{R}^3} |\hat{p}_{t_1,t_2}(u)| I(|u| > t^\alpha) du = O(t^{2d_X-1-\alpha}).$$

Then it remains to show that for any $0 < \alpha < \min(d_X/7, (1-2d_X)/8)$,

$$\sup_{|u|\le t^\alpha} |\hat{p}_{t_1,t_2}(u)| = O(t^{2d_X-1-4\alpha}), \tag{45}$$

as

$$\int_{\mathbb{R}^3} e^{-\mathbf{i}x\cdot u} \hat{p}_{t_1,t_2}(u_1,u_2,u_3) I(|u| \le t^\alpha) du \le C t^{3\alpha} t^{2d_X-1-4\alpha} = O(t^{2d_X-1-\alpha}).$$

To prove (45), write

$$\begin{aligned} \hat{f}_{X,t_1,t_2}(u_1,u_2,u_3) &= \prod_i \phi(u_1 b_{-i} + u_2 b_{t_1-i} + u_3 b_{t_2-i}) = \prod_{I_1}\cdots\prod_{I_2}\cdots\prod_{I_3}\cdots\prod_{I_4}\cdots \\ &=: a_1 \cdot a_2 \cdot a_3 \cdot a_4, \quad \text{say}, \\ \hat{f}_X(u_1)\hat{f}_{X,\Delta_t}(u_2,u_3) &= \prod_i \phi(u_1 b_{-i})\phi(u_2 b_{t_1-i} + u_3 b_{t_2-i}) = \prod_{I_1}\cdots\prod_{I_2}\cdots\prod_{I_3}\cdots\prod_{I_4}\cdots \\ &=: a'_1 \cdot a'_2 \cdot a'_3 \cdot a'_4, \quad \text{say}, \end{aligned}$$

where

$$\begin{aligned} I_1 &= \{i \in Z : |i| \le t^{2\alpha}\}, \quad I_2 = \{i \in Z : |i - t_1| \le t^{2\alpha}\}, \\ I_3 &= \{i \in Z : |i - t_2| \le t^{2\alpha}\}, \quad I_4 = Z \setminus (I_1 \cup I_2 \cup I_3). \end{aligned}$$

Then,

$$\begin{aligned} \hat{f}_{X,t_1,t_2}(u_1,u_2,u_3) - \hat{f}_X(u_1)\hat{f}_{X,\Delta_t}(u_2,u_3) &= a_1a_2a_3a_4 - a'_1a'_2a'_3a'_4 \\ &= (a_1 - a'_1)a_2a_3a_4 + a'_1(a_2 - a'_2)a_3a_4 \\ &\quad + a'_1a'_2(a_3 - a'_3)a_4 + a'_1a'_2a'_3(a_4 - a'_4). \end{aligned}$$

Hence, (45) will follow from the facts

$$a_i - a'_i = O(t^{2d_X-1-4\alpha}), \quad i = 1, 2, 3, \tag{46}$$

$$a_4 - a'_4 = -a'_4 u_1 u_2 \gamma(t_1) - a'_4 u_1 u_3 \gamma(t_2) + O(t^{2d_X-1-4\alpha}), \tag{47}$$

uniformly in $|u| \le t^\alpha$.

Using the inequality $|\prod c_i - \prod c_i'| \leq \sum |c_i - c_i'|$, for $|c_i| \leq 1$, $|c_i'| \leq 1$, one obtains

$$|a_1 - a_1'| \leq \sum_{|i| \leq t^{2\alpha}} |\phi(u_1 b_{-i} + u_2 b_{t_1 - i} + u_3 b_{t_2 - i}) - \phi(u_1 b_{-i})\phi(u_2 b_{t_1 - i} + u_3 b_{t_2 - i})|.$$

Moreover, for $|u_2| \leq t^\alpha$, $|u_3| \leq t^\alpha$, $|i| \leq t^{2\alpha}$, and $0 < \alpha < d_X/7$,

$$|u_2 b_{t_1 - i} + u_3 b_{t_2 - i}| \leq C t^{2d_X - 1 - 6\alpha}.$$

Therefore, as $|\phi(y_1 + y_2) - \phi(y_1)\phi(y_2)| \leq C|y_2|$, we obtain

$$|a_1 - a_1'| = O\Big(\sum_{|i| \leq t^{2\alpha}} t^{2d_X - 1 - 6\alpha}\Big) = O(t^{2d_X - 1 - 4\alpha}).$$

The cases $i = 2, 3$ are analogous.

It remains to prove (47). For sufficiently large t_1, t_2, the left-hand side of (47) can be represented as

$$a_4 - a_4' = a_4'(\exp\{Q_{t_1,t_2}(u_1, u_2, u_3)\} - 1),$$

where

$$Q_{t_1,t_2}(u_1, u_2, u_3) := \sum_{I_4} \psi(u_1 b_{-i}, u_2 b_{t_1 - i} + u_3 b_{t_2 - i})$$

with

$$\psi(x, y) := \log \phi(x + y) - \log \phi(x) - \log \phi(y), \quad (x, y) \in R^2.$$

Hence,

$$\begin{aligned}
&\psi(u_1 b_{-i}, u_2 b_{t_1 - i} + u_3 b_{t_2 - i}) \\
&= \int_0^{u_1 b_{-i}} \int_0^{u_2 b_{t_1 - i} + u_3 b_{t_2 - i}} \big[(\log \phi)''(0) + (x + y)(\log \phi)^{(3)}(z)\big] dx dy \\
&= -u_1 u_2 b_{-i} b_{t_1 - i} - u_1 u_3 b_{-i} b_{t_2 - i} + O\Big((u_1 b_{-i})^2 |u_2 b_{t_1 - i} + u_3 b_{t_2 - i}| \\
&\quad + |u_1 b_{-i}|(u_2 b_{t_1 - i} + u_3 b_{t_2 - i})^2\Big),
\end{aligned}$$

where z is between 0 and $x + y$. Consequently,

$$Q_{t_1,t_2}(u_1, u_2, u_3) = -u_1 u_2 \gamma_X(t_1) - u_1 u_3 \gamma_X(t_2) + q_1 + q_2,$$

where

$$q_1 = O\Big(|u_1u_2| \sum_{I_1\cup I_2\cup I_3} |b_{-i}b_{t_1-i}| + |u_1u_3| \sum_{I_1\cup I_2\cup I_3} |b_{-i}b_{t_2-i}|\Big),$$

$$q_2 = O\Big(|u_1^2u_2| \sum_{i=1}^{n} |b_{-i}^2 b_{t_1-i}| + |u_1^2u_3| \sum_{i=1}^{n} |b_{-i}^2 b_{t_2-i}| + |u_1u_2^2| \sum_{i=1}^{n} |b_{-i} b_{t_1-i}^2| + |u_1u_3^2| \sum_{i=1}^{n} |b_{-i} b_{t_2-i}^2|\Big).$$

As $|u_i| \le t^\alpha$, $i = 1, 2, 3$,

$$q_1 = O\Big(t^{2\alpha} \max_{|i|\le t^\alpha}(|b_{t_1-i}|, |b_{t_2-i}|)\Big) = O(t^{2d_X-1-5\alpha}),$$

$$q_2 = O\Big(t^{3\alpha} \max_{i}(|b_{t_1-i}|, |b_{t_2-i}|)\Big) = O(t^{2d_X-1-4\alpha}),$$

provided $0 < \alpha < d_X/7$. Thus

$$Q_{t_1,t_2}(u_1, u_2, u_3) = -u_1u_2\gamma_X(t_1) - u_1u_3\gamma_X(t_2) + O(t^{2d_X-1-4\alpha}) = O(t^{2d_X-1+2\alpha}),$$

uniformly in $]u_1|, |u_2|, |u_3| \le t^\alpha$, which implies that

$$\begin{aligned} \exp\{Q_{t_1,t_2}(u_1, u_2, u_3)\} - 1 &= Q_{t_1,t_2}(u_1, u_2, u_3) + O\big(Q^2_{t_1,t_2}(u_1, u_2, u_3)\big) \\ &= -u_1u_2\gamma_X(t_1) - u_1u_3\gamma_X(t_2) + O(t^{2d_X-1-4\alpha}) \end{aligned}$$

as $\alpha < (1 - 2d_X)/8$. This proves (47) and the lemma too.

References

1. Csörgo, S., Mielniczuk, J.: Nonparametric regression under long-range dependent normal errors. Ann. Statist. **23**, 1000–1014 (1995)
2. Dalla, V., Giraitis, L., Hidalgo, J.: Consistent estimation of the memory parameter for nonlinear time series. J. Time Ser. Anal. **27**, 211–251 (2006)
3. Fernholz, L.: Almost sure convergence of smoothed empirical distribution functions. Scand. J. Statist. **18**, 255–262 (1991)
4. Giraitis, L., Koul, H.L., Surgailis, D.: Asymptotic normality of regression estimators with long memory errors. Statist. Probab. Lett. **29**, 317–335 (1996)
5. Giraitis, L., Koul, H.L., Surgailis, D.: Large sample inference for long memory processes. Imperial College Press, London (2012)
6. Guo, H., Koul, H.: Asymptotic inference in some heteroscedastic regression models with long memory design and errors. Ann. Statist. **36**, 458–487 (2008)
7. Koul, H.L., Mimoto, N., Surgailis, D.: Goodness-of-fit tests for long memory moving average marginal density. Metrika **76**, 205–224 (2013)

8. Koul, H.L., Surgailis, D.: Asymptotic expansion of the empirical process of long memory moving averages. In: Dehling, H., Mikosch, T., Sørensen, M. (eds.) Empirical Process Techniques for Dependent Data, pp. 213–239. Birkhäser, Boston (2002)
9. Koul, H.L., Surgailis, D.: Goodness-of-fit testing under long memory. J. Statist. Plann. Inference **140**, 3742–3753 (2010)
10. Koul, H.L., Surgailis, D.: Goodness-of-fit tests for marginal distribution of linear random fields with long memory. RM # 702, Department of Statistics & Probability, Michigan State University (2013)
11. Kulik, R., Wichelhaus, C.: Nonparametric conditional variance and error density estimation in regression models with dependent errors and predictors. Electron. J. Stat. **5**, 856–898 (2011)
12. Liu, R., Yang, L.: Kernel estimation of multivariate cumulative distribution function. J. Nonparametr. Stat. **20**, 661–677 (2008)
13. Lorek, P., Kulik, R.: Empirical process of residuals for regression models with long memory errors. Statist. Probab. Lett. **86**, 7–16 (2014)
14. Masry, E., Mielniczuk, J.: Local linear regression estimation for time series with long-range dependence. Stochastic Process. Appl. **82**, 173–193 (1999)
15. Müller, U.U., Schick, A., Wefelmeyer, W.: Estimating the innovation distribution in nonparametric autoregression. Probab. Theory Relat. Fields **144**, 53–77 (2009)
16. Robinson, P.M.: Semiparametric analysis of long memory time series. Ann. Statist. **22**, 515–539 (1994)
17. Robinson, P.M.: Gaussian semiparametric estimation of long range dependence. Ann. Statist. **23**, 1630–1661 (1995)
18. Robinson, P.M.: Large-sample inference for nonparametric regression with dependent errors. Ann. Statist. **25**, 2054–2083 (1997)
19. Robinson, P.M.: Asymptotic theory for nonparametric regression with spatial data. J. Econometrics **165**, 5–19 (2011)
20. Wang, J., Liu, R.: Cheng, F., Yang, L.: Oracally efficient estimation of autoregressive error distribution with simultaneous confidence band. Ann. Statist. **42**, 654–668 (2014)
21. Zygmund, A.: Trigonometric series, 3rd edn. Cambridge University Press, Cambridge (2002)

Testing Shape Constraints in Lasso Regularized Joinpoint Regression

Matúš Maciak

Abstract Joinpoint regression models are very popular in some practical areas mainly due to a very simple interpretation which they offer. In some situations, moreover, it turns out to be also useful to require some additional qualitative properties for the final model to be satisfied. Especially properties related to the monotonicity of the fit are of the main interest in this paper. We propose a LASSO regularized approach to estimate these types of models where the optional shape constraints can be implemented in a straightforward way as a set of linear inequalities and they are considered simultaneously within the estimation process itself. As the main result we derive a testing approach which can be effectively used to statistically verify the validity of the imposed shape restrictions in piecewise linear continuous models. We also investigate some finite sample properties via a simulation study.

Keywords Joinpoint regression · Regularization · Shape constraints · Post-selection inference

1 Introduction

There are many different areas and practical situations where an easy interpretation of some statistical model is of the key interest. This especially happens if statistical models are about to be used by nonstatisticians or people with a lack of statistical skills. If possible, in such cases, statisticians should have a tendency to go for models which are flexible enough to capture an assumed underlying structure behind the data, but, on the other hand, the final models should be simple enough to be easily explained and interpreted.

M. Maciak (✉)
Department of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic
e-mail: maciak@karlin.mff.cuni.cz

J. Antoch et al. (eds.), *Analytical Methods in Statistics*, Springer Proceedings in Mathematics & Statistics 193, DOI 10.1007/978-3-319-51313-3_6

Joinpoint regression[1] models definitely belong among this type of models. Each joinpoint regression model is a piecewise linear model where each single piece can be interpreted in a classical sense of an ordinary linear regression line. The linear pieces join together obeying the continuity condition over the whole domain of interest. Locations at which the linear segments join together are called change-points (sometimes also structural breaks, join points, slope breaks or transition points). The locations of these change-points are usually left unknown and they are also considered to be the subjects of the estimation process.

The joinpoint regression models take an advantage of changes in the slope parameter to adapt for existing alternations in the overall structure in the data. However, in order to keep the final model simple, it is usually assumed that only some change-points are present in the model and thus, only some breaks in the slope are observed in the final fit. From the theoretical point of view it turns out to be convenient to know the number of change-points in advance as this knowledge improves the asymptotic performance of the model. From practical reasons however, as we already mentioned, this is not the case in real situations nor it is possible many times. On the other hand, if the number of change-points (slope breaks) is unknown one needs to use some model selection procedures to perform a model selection step: a proper decision needs to be made in order to choose one model from a set of all allowable models. There are of course various techniques to perform this model selection step: considering a selection of a final model from some class of plausible joinpoint linear regression models (a set of models with different number of breaks in the slope, or equivalently, models with the corresponding number of linear segments) one can either use classical statistical tests (e.g. permutation tests investigated in [10, 11] or likelihood ratio tests discussed in [4, 8]) or instead some alternative approaches mostly based on a Bayesian framework (see, e.g. [2] or [15]). In this context, there are especially permutations tests gaining a lot of popularity (see [1, 19]) as they are commonly used as standard tools in different areas where joinpoint models are applied.

Another idea to perform the model selection step, which we also use in this paper, is based on the LASSO regularization approach originally proposed in [20]. The same idea of using the LASSO approach in joinpoint regression estimation is investigated in [14] and also in [7] where, however, the authors considered a slightly different but still very analogous scenario with a locally constant model which represents the simplest structure one can consider in this setup. The main idea behind the LASSO-based selection step is to use the L_1 norm penalty implemented in the minimization of some objective function and to let the penalty itself to select one model from a set of all plausible ones. To put this in other words, the L_1 penalty is responsible for choosing only important breaks (change-points) in the slope from some much larger set of all hypothetical slope breaks. Only a small subset of important change-points should play the role in the final fit—the rest of the change-points will stay inactive.

[1]Different names can be used in literature to refer to joinpoint regression models, among others, for instance segmented regression models, piecewise linear models, threshold models, sequential linear models, etc.

The solution is called to be sparse (i.e. only a small subset of all corresponding parameters are nonzero in the final model) and creating sparse solutions is the main property of the LASSO regularized approach (see, e.g. [20] for more details). The main advantage of using the LASSO approach is that the regularized estimation somehow considers all possible model alternatives and the L_1 penalty selects the most suitable model in a data-driven manner. In contrast, the permutation tests and Bayesian approaches as well are quite limited in this context as they mostly consider only some small number of different alternatives. In permutation tests one usually tests one specific model against some other one and in the Bayesian framework the prior is usually also defined over a small set of different models only. From this point of view the LASSO-based selection step seems to be much more appropriate strategy especially if the number of slope breaks in the true model is unknown (see also [14] for further discussion).

The main idea of this paper is to derive a testing approach to verify some additional qualitative constrains which might be (optionally) imposed on the shape of the constructed fit. As far as we assume the piecewise linear models only the most reasonable restrictions which can be formulated with respect to the overall shape are those related to monotonic properties. One can, for example, requires that the final fit should be increasing or non-decreasing over the whole domain of interest. Such models have practical applications especially in production function estimation, different cost and performance functions, but they are also useful in economics, medicine, social sciences and many other areas.

In addition to the tests of qualitative properties the designed testing approach can be also applied in a straightforward way to verify statistical importance of change-points being estimated in the model.

The rest of the paper is organized as follows: in the next session we discuss the LASSO regularized joinpoint models and we propose an algorithm to estimate these models. Different shape constraints which can be automatically considered in the estimation procedure are also discussed in Sect. 2. In Sect. 3 we discuss the main result of the paper: the testing approach based on some modifications of some recent results in [21]. All necessary details are also provided. Finally, finite sample properties and applications are presented and discussed in Sect. 4.

2 Joinpoint Regression Model with Shape Constrains

Let us consider a bivariate random sample $\{(X_i, Y_i);\ i = 1, \ldots, n\}$ drawn from some unknown population (X, Y), where the design points $X_1, \ldots, X_n$ are assumed to be drawn from some continuous marginal distribution from some domain of interest, e.g. interval $[0, 1]$. Thus, all X_i points are unique almost surely and with any loss of generality we can assume that $X_i < X_{i+1}$, for all $i = 1, \ldots, n-1$. The unknown dependence structure between each pair of the random variables Y_i's and X_i's is assumed to be piecewise linear and continuous, and the unknown linear pieces are allowed to join at the design points $X_1, \ldots, X_n$ only. Under this assumptions a general joinpoint regression model with no further shape restrictions can be expressed as

$$Y_i = a_i + b_i X_i + \varepsilon_i, \qquad \text{for } i = 1, \ldots, n, \tag{1}$$

where we assume independent random error terms $\varepsilon_i \sim N(0, \sigma^2)$, for some given constant $\sigma^2 > 0$. The overall continuity condition can be imposed as

$$a_i + b_i X_i = a_{i+1} + b_{i+1} X_i, \qquad \text{for } i = 1, \ldots, n-1. \tag{2}$$

Using this model formulation we are firstly interested in estimating the unknown parameters $a_i, b_i \in \mathbb{R}$, for $i = 1, \ldots, n$. However, as far as we only assume some small number of slope breaks to occur in the final model we can consider parameters b_i's to be sparse in a sense that equalities $b_i = b_{i+1}$ hold for all but some small subset of indexes from $\{1, \ldots, n-1\}$. Therefore, it sounds reasonable to apply the estimation idea based on the LASSO regularization principle where the parameters included in the L_1 penalty are the corresponding differences between two neighbouring slope parameters b_i and b_{i+1}, for $i = 1, \ldots, n-1$.

Once there is some index j such that $b_j \neq b_{j+1}$ the estimation procedure automatically estimates a slope change in the model and the corresponding location of the slope break is estimated at the design point X_j. The overall slope changes at this location from b_j to b_{j+1} to adjust for an existing structural change in the data. The intercept parameters a_i, for $i = 1, \ldots, n$, are then responsible for obeying the continuity condition over the whole domain.

Let $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \ldots, \beta_{n-1})^\top = (a_1, b_1, (b_2 - b_1), \ldots, (b_{n-1} - b_{n-2}))^\top \in \mathbb{R}^n$ and $\mathscr{I} = \{2, \ldots n-1\}$ the subset of indexes from $\{0, 1, \ldots, n-1\}$ for which the vector of parameters $\boldsymbol{\beta}$ is assumed to be sparse. Then, using some simple algebra calculations one can easily verify that an appropriate estimation procedure for the model in (1) under the continuity conditions in (2) can be formulated as

$$\underset{\boldsymbol{\beta} \in \mathbb{R}^n}{\text{Minimize}} \frac{1}{n} \|\boldsymbol{Y} - \mathbb{X}\boldsymbol{\beta}\|_2^2 + \lambda_n \left\|\boldsymbol{\beta}_{(-2)}\right\|_1, \tag{3}$$

where $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^\top$ is the response vector, $\boldsymbol{\beta} \in \mathbb{R}^n$ is the vector of unknown parameters to be estimated and $\boldsymbol{\beta}_{(-2)} = (\beta_2, \ldots, \beta_{n-1})^\top$ denotes the sparse subset of parameters (i.e. the whole parameter vector without its first two elements) which are regularized by the LASSO penalty in (3). If there are some non-zero elements estimated for $\boldsymbol{\beta}_{(-2)}$ these estimates are responsible for slope breaks, respectively, changes in the direction of the dependence of Y on X. One can also easily verify that the design matrix in (3) takes the form

$$\mathbb{X} = \begin{pmatrix} 1 & X_1 & 0 & 0 & \dots & 0 \\ 1 & X_2 & 0 & 0 & \dots & 0 \\ 1 & X_3 & (X_3 - X_2) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & X_n & (X_3 - X_2) & (X_4 - X_3) & \dots & (X_n - X_{n-1}) \end{pmatrix}. \tag{4}$$

The first two columns in $\mathbb{X}$ are identical with the columns used in the design matrix for a classical linear regression case where Y is regressed on X. Thus, the first two parameters can be interpreted as the overall intercept and slope parameters. All other parameter estimates, where only some of them are expected to be nonzero due to the L_1 penalty in (3), are responsible for estimating existing slope changes. The value of the regularization parameter $\lambda_n > 0$ in (3) controls the number of these slope changes in the final model: for λ_n close to zero there should be slope changes expected at each X_i while for $\lambda_n \to \infty$ (respectively, large enough) the final fit will perfectly correspond with a classical linear regression line.

The minimization problem in (3) is convex and thus it can be easily solved using some standard minimization toolboxes. Moreover, one can also take an advantage of very effective solution approaches like LARS–LASSO algorithm (see [3]) or the coordinate decent algorithm proposed in [5]. The whole minimization problem in (3) can be also reparametrized in a sense of the classical LASSO problem (see [14]). More details about the model in (1) including some theoretical properties of the estimates constructed by minimizing (3) are discussed in [14].

Next, we would like to point our attention to situations where it may be natural to expect some specific shape of the unknown dependence structure. In order to obtain the same structure in the final estimate one needs to impose some additional constraints in order to enforce the required shape. As far as we only deal with the dependence structures which are piecewise linear it is reasonable to mainly consider restrictions based on monotonicity. One can still assume some isotonic constraints but let us remind that with piecewise modelling approach it is not possible to obtain any strictly convex or concave shapes.

Using the notation already introduced above, it is easy to see, for example, that for an (strictly) increasing form of the dependence between Y and X it needs to hold that $b_i \geq 0$ (or $b_i > 0$, respectively.) for each $i \in \{1, \dots, n\}$. Considering the vector of parameters $\boldsymbol{\beta} \in \mathbb{R}^n$ it gives that $\sum_{j=1}^{\ell} \beta_j \geq 0$ (or $\sum_{j=1}^{\ell} \beta_j > 0$, respectively.) which needs to hold for any $\ell \in \{1, \dots, n-1\}$. Analogously, if the underlying shape of the function is supposed to be (strictly) decreasing the same inequalities should hold with opposite signs. For a convex shape of the piecewise linear dependence one can again verify that it needs to hold that $b_{i+1} \geq b_i$ for all $i \in \{1, \dots, n-1\}$ which again can be equivalently reformulated in terms of the vector of parameters $\boldsymbol{\beta} \in \mathbb{R}^n$ as $\beta_j \geq 0$ for each $j \in \{2, \dots, n-1\}$. Similarly, for the concave shape we need opposite inequalities to be satisfied. Note that no role is indeed played by the overall intercept parameter β_1 in case of isotonic restrictions, and thus, this parameter is not present in the corresponding constraints. In some situations it may be reasonable to even combine the monotonic and isotonic restrictions together. This can be done in a straightforward way using the inequalities already defined.

In order to obtain an estimate of the unknown dependence structure between Y and X which will in addition fully comply with the imposed shape constraints, the corresponding inequalities need to be taken into account simultaneously with the estimation process. Again, it is easy to see that in order to fit a LASSO regularized shape-constrained joinpoint regression model one needs to consider the minimization problem which can be expressed as

$$\underset{\boldsymbol{\beta} \in \mathbb{R}^n}{\text{Minimize}} \frac{1}{n} \|\boldsymbol{Y} - \mathbb{X}\boldsymbol{\beta}\|_2^2 + \lambda_n \left\|\boldsymbol{\beta}_{(-2)}\right\|_1, \tag{5}$$

$$\text{subject to } \boldsymbol{A}\boldsymbol{\beta}_{(-1)} \geq 0, \tag{6}$$

for some appropriate matrix[2] $\boldsymbol{A}$, where in addition $\boldsymbol{\beta}_{(-1)} \in \mathbb{R}^{n-1}$ denotes analogously a subset of $\boldsymbol{\beta}$ given by the vector itself without its first element (the overall intercept parameter). The inequality in (6) is meant elementwise.

The minimization problem under the constraints given in (5) and (6) looks very similar to the CLASSO problem proposed in [6] where a classical LASSO problem with additional linear constraints imposed on estimated parameters is considered. However, unlike the CLASSO problem we have different subsets of parameters playing now their role in different parts of the minimization problem: the whole set of parameters (vector $\boldsymbol{\beta} \in \mathbb{R}^n$) contributes to the minimization of the L_2 loss criterion function in (5) while only a subset of parameters (i.e. vector $\boldsymbol{\beta}_{(-2)} \in \mathbb{R}^{n-2}$) is considered in the L_1 penalty term. Finally, in the shape restriction imposed via inequalities in (6) another subset of parameters is used to define the required shape—either parameters in $\boldsymbol{\beta}_{(-1)} \in \mathbb{R}^{n-1}$ in case of the monotonic restrictions or parameters in $\boldsymbol{\beta}_{(-2)} \in \mathbb{R}^{n-2}$ in case of the isotonic restrictions.

Unfortunately, the minimization problem under the linear constraints defined in (5) and (6) can not be re-parametrized to fit the classical LASSO problem as we could do it in case of the unrestricted version stated in (3). The minimization problem which we consider for the regularized and shape-constrained joinpoint models does

[2] If the shape constraints in (6) refer to the monotonic property of the final fit (e.g.non-decreasing function), the corresponding matrix $\boldsymbol{A}$ equals to $\boldsymbol{A}_1$ below. If the constraints in (6) are supposed to refer to isotonic property (e.g. convex function) the corresponding matrix $\boldsymbol{A}$ should be equal to $\boldsymbol{A}_2$ (the estimate for the overall slope β_1 is irrelevant for isotonic properties of the final fit, thus the first line in $\boldsymbol{A}_2$ are either zeros or it can be deleted with $\boldsymbol{\beta}_{(-2)}$ being considered instead of $\boldsymbol{\beta}_{(-1)}$).
$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \dots & 1 & 1 & 0 \\ 1 & \dots & \dots & 1 & 1 \end{pmatrix} \in \mathbb{R}^{(n-1)\times(n-1)}$ and $\boldsymbol{A}_2 = \begin{pmatrix} 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix} \in \mathbb{R}^{(n-1)\times(n-1)}$. Analogous matrices for other scenarios can be obtained in similar ways as an easy exercise.

not fit any standard LASSO problems considered in literature therefore, it needs to be considered separately.

Moreover, one can not even apply classical LASSO estimation tools like LARS–LASSO approach or coordinate decent algorithm (see the discussion, e.g. in [6]). On the other hand, the shape restricted LASSO minimization problem as we consider it in this paper is still a convex problem and thus, it can be solved using some standard optimization toolboxes and iterative algorithms.[3]

3 Statistical Test for Testing Shape Constraints Validity

In this section we propose a suitable testing approach which can be used to decide whether some shape constraints considered for the unknown joinpoint model are statistically relevant or not. To be precise, the main idea of this section is to provide a statistical test to decide about the following pair of hypothesis

$$\begin{aligned} H_0: &\quad \boldsymbol{A}\boldsymbol{\beta}_{(-1)} \geq 0; \\ H_1: &\quad H_0 \quad does\ not\ hold; \end{aligned} \tag{7}$$

where the inequality is again assumed in the element-wise manner and the matrix $\boldsymbol{A}$ in the null hypothesis is some appropriate matrix with the corresponding dimensions: it is used to expresses the required shape constraints in terms of some linear combination of the unknown parameters. Note that any general matrix form is allowable for the testing procedure discussed in this section, however, we only concentrate on such forms which have some straightforward interpretation with respect to monotonic or isotonic properties of the LASSO regularized joinpoint regression model.

Considering an analogous set of hypotheses in a classical linear regression model one can use a whole variety of different statistical tools to decide whether the null hypothesis in (7) should be rejected or not. Indeed, one can use for example a classical test based on estimated residuals in the model which corresponds with H_0 and the other model corresponding with H_1 and to compare both. The test statistics then follows Fisher F-distribution with some appropriate degrees of freedom. Another option is to go for a likelihood ratio test and again we need to compare two models against each other. The corresponding test statistics then follows χ^2 distribution with some degrees of freedom. Last but not least, one can also use a series of tests where for each element-wise inequality in (7) a separate t-test is performed, however, with some additional correction in order to keep a predefined level of the first type error probability. Unfortunately, none of these approaches can be used in case of the LASSO regularized estimates as one can not guarantee the overall first type error probability any more.

---

[3]In order to solve the minimization problem in (6) with respect to the constraints stated in (7) we use The MOSEK optimization toolbox and Mosek-to-R interface available in the R package `Rmosek`.

Let us briefly discuss some main issues and problems which commonly arise when performing statistical inference about LASSO regularized estimates in general. The same issues also occur for the shape constraint test in (7) as the corresponding estimates are also obtained by the LASSO regularized minimization in (6) and (7). We also discuss some recent proposals which can avoid some of these problems and finally, we introduce an adaptation which perfectly suits our scenario: we propose a test which can be effectively used to verify whether the imposed shape constraints are statistically relevant in the joinpoint model or not.

3.1 Significancy Test for LASSO Regularized Estimates

There is a huge discussion in literature on different approaches on how to perform statistical inference and statistical tests about some parameters estimated within the LASSO regularized regression framework. A pioneering work illustrating a whole series of problems arising when testing LASSO regularized parameters is discussed in [13].

The whole drawback is hidden in the fact that in the LASSO regularized estimation process the non-zero parameters enter the model at random. In other words, selecting a covariate, respectively, a column from the design matrix $\mathbb{X}$ in (3), is a random event and all classical testing approaches are not able to take this into account. This results in a poor performance with respect to the ability to keep the predefined level of the test (see [13] for an illustration).

There is a recent and extensive development in the area of the post-selection inference and a whole series of different approaches are proposed: e.g. methods based on splitting the sample and re-sampling discussed in [16] or ideas derived from debiasing the LASSO estimates proposed by [23] or [22]. In this paper we use the idea presented in [21] where the authors showed that the problem of random selection in LASSO models can be effectively taken care of by conditioning on the actual LASSO path history.[4] In addition, this LASSO path history can be expressed in terms of some well-defined linear constraints which define a polyhedral set as a subset of $\mathbb{R}^n$ for the vector of observations $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$, which gives the same LASSO history path. Thus, instead of considering all possible data scenarios $\boldsymbol{Y} \in \mathbb{R}^n$ we only consider those which give the same LASSO solution path. We condition on $\boldsymbol{Y} \in \mathbb{B}$ where $\mathbb{B} \subset \mathbb{R}^n$ is some well-defined polyhedral set which fully defines the actual LASSO history path.

Thus, considering a set of LASSO regularized parameters $\boldsymbol{\beta}_{(-2)} \in \mathbb{R}^{n-2}$ and some general statistical test

[4] By LASSO path history we understand a sequence of variables which progressively enter the model as we proceed with estimation. In each step of the estimation procedure only one parameter can either enter the current model (if it is not in the model yet) or an active parameter steps off the model (if it was active). Only one of these events can happen at each step. The LASSO history follows this entering/stepping off process starting with a zero model where all regularized parameters are set to zero.

$$\begin{aligned} H_0 &: \boldsymbol{v}^\top \boldsymbol{\beta}_{(-2)} = 0; \\ H_1 &: \boldsymbol{v}^\top \boldsymbol{\beta}_{(-2)} < 0; \end{aligned} \tag{8}$$

for some arbitrary vector $\boldsymbol{v} \in \mathbb{R}^{n-2}$, we can apply the aforementioned theory based on conditioning on the LASSO history path to replace the conditional probability under the null hypothesis

$$P_{H_0}(\cdot|\ full\ LASSO\ history\ path)$$

by a more appropriate (mathematical) expression

$$P_{H_0}(\cdot|\boldsymbol{Y} \in \mathbb{B}) = P_{H_0}(\cdot|\{\boldsymbol{Y} \in \mathbb{R}^n;\ \varGamma\boldsymbol{Y} \geq \boldsymbol{b}\}), \tag{9}$$

for some appropriate polyhedral set $\mathbb{B} = \{\boldsymbol{Y} \in \mathbb{R}^n;\ \varGamma\boldsymbol{Y} \geq \boldsymbol{b}\} \in \mathbb{R}^n$ which fully specifies the LASSO history path in terms of some linear constraints imposed on $\boldsymbol{Y}$ being expressed as $\varGamma\boldsymbol{Y} \geq \boldsymbol{b}$. Finally, an appropriate test statistic denoted as $T(\boldsymbol{Y}, \mathbb{B}, \boldsymbol{v})$ can be derived for the set of hypothesis in (8) and it is proved in [21] that under the null hypothesis it holds that

$$T(\boldsymbol{Y}, \mathbb{B}, \boldsymbol{v}) \overset{P_{H_0}}{\sim} Uniform(0, 1).$$

We use this result in the next session to propose a modification which we later use to decide on statistical significancy of the proposed shape constraints in the LASSO regularized joinpoint regression model.

3.2 Shape Constraints Inference for LASSO Joinpoint Models

Let us start with some technical details on how to transform the full LASSO history path into a polyhedral expression which is used for the conditional probability in (9). This will be needed to calculate the distribution of the test statistic under the null hypothesis discussed later.

Let $\boldsymbol{X}_j$ for $j = 1, \ldots, n$ denotes the corresponding column of the design matrix $\mathbb{X}$. Then for some vector of coefficient estimates $\widehat{\boldsymbol{\beta}} = (\beta_0, \ldots, \beta_{n-1})^\top \in \mathbb{R}^n$ and the vector of signs[5] $\hat{\boldsymbol{s}} = (s_1, \ldots, s_{n-2})^\top \in [-1, 1]^{(n-2)}$ to be the solution of (5) with additional constraints (6) it is necessary and sufficient than the Karush–Kuhn–Tucker (KKT) optimality conditions hold. They can be expressed as

[5]The vector of estimated signs $\hat{\boldsymbol{s}}$ is only relevant for the LASSO regularized estimates in $\boldsymbol{\beta}_{(-2)}$ and it holds that $\hat{s}_j = sign(\hat{\beta}_{j+1})$, for all $j = 1, \ldots, n-2$.

$$
\begin{array}{ll}
\boldsymbol{X}_j^\top(\mathbb{X}\widehat{\boldsymbol{\beta}} - \boldsymbol{Y}) = 0 & \text{for } j = 1, 2, \\
\boldsymbol{X}_j^\top(\mathbb{X}\widehat{\boldsymbol{\beta}} - \boldsymbol{Y}) + \lambda_n \hat{s}_{j-2} = 0 & \text{for } j = 3, \ldots, n, \\
\hat{s}_i = sign(\hat{\beta}_{i+1}) & \text{if } \hat{\beta}_{i+1} \neq 0 \text{ for } i = 1, \ldots, n-2, \\
\hat{s}_i \in [-1, 1] & \text{if } \hat{\beta}_{i+1} = 0,
\end{array}
$$

where in addition we require the shape constraints to hold, thus $\boldsymbol{A}\widehat{\boldsymbol{\beta}}_{(-1)} \geq 0$ again for $\widehat{\boldsymbol{\beta}}_{(-1)} = (\widehat{\beta}_1, \ldots, \widehat{\beta}_{n-1})^\top$. Using now Lemma 4.1 and Theorem 4.3 from [12] we can use the KKT conditions above to derive the LASSO selection history in terms of some affine restrictions imposed on $\boldsymbol{Y} \in \mathbb{R}^n$. Thus, a random event that we obtain a model with the same active set $\mathscr{A} = \{j \in \mathscr{J};\ \widehat{\beta}_j \neq 0\}$ and the same vector of corresponding signs $\hat{\boldsymbol{s}}$, where $\hat{s}_j = sign(\hat{\beta}_{j+1})$, for $j = 1, \ldots, n-2$, can be expressed as in terms of a polyhedral restriction on $\boldsymbol{Y} \in \mathbb{R}^n$, respectively

$$
\mathbb{B} = \{\boldsymbol{Y} \in \mathbb{R}^n | \mathscr{A}, \hat{\boldsymbol{s}}\} = \{\boldsymbol{Y} \in \mathbb{R}^n;\ \varGamma\boldsymbol{Y} \geq \boldsymbol{b}\},
$$

where both, matrix $\varGamma$ and vector $\boldsymbol{b}$ depends on the LASSO selection history given by the active set of indexes $\mathscr{A}$ and the corresponding sign vector $\hat{\boldsymbol{s}}$.

Next, we state Lemma 1 from [21] which will be useful to define the test statistic for the test.

Lemma 1 *Let $\boldsymbol{Y} \sim N(\boldsymbol{\theta}, \varSigma)$ for some general mean vector $\boldsymbol{\theta} \in \mathbb{R}^n$ and variance–covariance matrix $\varSigma$ such that $\boldsymbol{v}^\top \varSigma \boldsymbol{v} \neq 0$ for an arbitrary fixed vector $\boldsymbol{v} \in \mathbb{R}^n$. Then a polyhedral restriction $\{\varGamma\boldsymbol{Y} \geq \boldsymbol{b}\}$ can be equivalently expressed as*

$$
V_L(\boldsymbol{Y}) \leq \boldsymbol{v}^\top \boldsymbol{Y} \leq V_U(\boldsymbol{Y}) \text{ and at the same time } V_0(\boldsymbol{Y}) \leq 0,
$$

where for $V_L(\boldsymbol{Y})$, $V_U(\boldsymbol{Y})$ and $V_0(\boldsymbol{Y})$ we have

$$
V_L(\boldsymbol{Y}) = \max_{j:\rho_j>0} \frac{u_j - (\varGamma\boldsymbol{Y})_j + \rho_j \boldsymbol{v}^\top V_L(\boldsymbol{Y})}{\rho_j}
$$

$$
V_U(\boldsymbol{Y}) = \max_{j:\rho_j<0} \frac{u_j - (\varGamma\boldsymbol{Y})_j + \rho_j \boldsymbol{v}^\top V_L(\boldsymbol{Y})}{\rho_j}
$$

and finally, $V_0(\boldsymbol{Y}) = \max_{j:\rho_j=0} u_j - (\varGamma\boldsymbol{Y})_j$, where $\boldsymbol{\rho} = (\rho_j)_j = \varGamma\varSigma\boldsymbol{v}/\boldsymbol{v}^\top\varSigma\boldsymbol{v}$. Moreover, $(V_L(\boldsymbol{Y}), V_U(\boldsymbol{Y}), V_0(\boldsymbol{Y}))$ is independent of $\boldsymbol{v}^\top\boldsymbol{Y}$.

Proof See the proof of Lemma 1 in [21]. □

Now, we only need to realize that $\boldsymbol{Y}$ follows by our assumption the normal distribution with the corresponding variance–covariance matrix $\varSigma$ which we assume is known. A linear combination $\boldsymbol{v}^\top\boldsymbol{Y}$ for an arbitrary $\boldsymbol{v} \in \mathbb{R}^n$ follows again the normal distribution and thus, a conditional distribution when conditioning on the polyhedral set $\{\Gamma \boldsymbol{Y} \geq \boldsymbol{b}\}$ can be equivalently expressed as a conditional distribution of

$$\boldsymbol{v}^\top \boldsymbol{Y} \;\Big|\; V_L(\boldsymbol{Y}) \leq \boldsymbol{v}^\top \boldsymbol{Y} \leq V_U(\boldsymbol{Y}) \bigwedge V_0(\boldsymbol{Y}) \leq 0,$$

which can be shown is a truncated normal distribution[6] where the boundaries of the truncation are random and depend on the selection.

Finally, to construct the test statistic which can be later used to test the set of hypotheses in (8) we state the following theorem.

Theorem 1 *Let the model* $\boldsymbol{Y} \sim N(\mathbb{X}\boldsymbol{\beta}, \sigma^2 I)$ *holds for some general mean vector* $\mathbb{X}\boldsymbol{\beta} \in \mathbb{R}^n$ *and some given variance* $\sigma^2 > 0$. *Let moreover,* $\boldsymbol{v} \in \mathbb{R}^n$ *be an arbitrary fixed vector such that* $\boldsymbol{v}^\top(\mathbb{X}^\top\mathbb{X})^{-1}\boldsymbol{v} \neq 0$. *Then the random variable*

$$T(\boldsymbol{Y}, \mathbb{B}, \boldsymbol{v}) = 1 - F^{[V_L(\boldsymbol{Y}), V_U(\boldsymbol{Y})]}_{\boldsymbol{v}^\top\boldsymbol{\beta}, \boldsymbol{v}^\top(\mathbb{X}^\top\mathbb{X})^{-1}\boldsymbol{v}}\left(\boldsymbol{v}^\top\widehat{\boldsymbol{\beta}}\right), \tag{10}$$

can be used to test the null hypothesis $H_0 : \boldsymbol{v}^\top\boldsymbol{\beta} = 0$ *against the alternative* $H_1 : \boldsymbol{v}^\top\boldsymbol{\beta} > 0$ *and it holds that*

$$P_{H_0}\Big(T(\boldsymbol{Y}, \mathbb{B}, \boldsymbol{v}) \leq \alpha | full\ LASSO\ history\ path\Big)$$
$$= P_{\boldsymbol{v}^\top\boldsymbol{\beta}=0}\Big(T(\boldsymbol{Y}, \mathbb{B}, \boldsymbol{v}) \leq \alpha | \Gamma\boldsymbol{Y} \geq \boldsymbol{b}\Big) = \alpha.$$

Proof The proof follows as a special case of the proof of Lemmas 2 and 3 in [21]. □

Comparing now the set of hypothesis in (7) with the post-selection tests discussed in the previous section one should note two major differences: there is only a simple null hypothesis considered in (8) but the formulation in (7) states a more complex hypothesis which can not be directly dealt with using the discussed post-selection inference above. And second, unlike the null hypothesis in (8) where only one equality is given one needs to consider a whole set of inequalities in the expression $\boldsymbol{A}\boldsymbol{\beta}_{(-1)} \geq 0$ in the null hypothesis in (7) as this inequality is considered in the element-wise manner. Indeed, there are actually $n - 1$ inequalities to test in total. Unfortunately, the post-selection tests can not be easily adapted for such complex hypotheses and therefore, some necessary modifications and ways around need to be investigated.

On the other hand, we apriori assume that only a small number of parameters in $\boldsymbol{\beta}_{(-2)} \in \mathbb{R}^{n-2}$ are estimated as non-zero elements and thus, we can equivalently rewrite the original test in (7) using only those parameters where non-zero elements occur in the estimate of $\boldsymbol{\beta}_{(-2)}$. In addition, we can also limit our attention to a more

[6]By a truncated normal distribution we understand a distribution with a cumulative distribution function $F^{[a,b]}_{\mu,\sigma^2}(x)$ truncated to the interval $[a, b] \subset \mathbb{R}$ where $F^{[a,b]}_{\mu,\sigma^2}(x) = \frac{\phi((x-\mu)/\sigma) - \phi((a-\mu)/\sigma)}{\phi((b-\mu)/\sigma) - \phi((a-\mu)/\sigma)}$, for $\phi(\cdot)$ being the cumulative distribution function of the standard Gaussian random variable.

specific scenario, where we only test a simple null hypothesis against a complex alternative. Given the fact that the testing approach is based on the post-selection inference where we condition on the full LASSO history path (which means that we condition on scenarios which give the same final model) we can rewrite the set of hypothesis in (7) using $\mathscr{A} = \{j \in \mathscr{J};\ \hat{\beta}_j \neq 0\}$, which is the set of indexes of regularized parameters which are active in the final fit, as

$$\begin{aligned} H_0 :\ & \boldsymbol{A}_{\mathscr{A}_{(1)}} \boldsymbol{\beta}_{\mathscr{A}_{(1)}} = 0; \\ H_1 :\ & \boldsymbol{A}_{\mathscr{A}_{(1)}} \boldsymbol{\beta}_{\mathscr{A}_{(1)}} > 0; \end{aligned} \tag{11}$$

where $\mathscr{A}_{(1)} = \mathscr{A} \cup \{j = 1\}$ to also include the intercept parameter β_1 into the test,[7] and $\boldsymbol{A}_{\mathscr{A}_{(1)}}$ is a submatrix of $\boldsymbol{A}$ consisting only of the rows and columns which correspond with the set of indexes in $\mathscr{A}_{(1)}$. The same also applies for $\boldsymbol{\beta}_{\mathscr{A}_{(1)}}$ which is a subvector of $\boldsymbol{\beta}_{(-1)} \in \mathbb{R}^{n-1}$ with only those elements which correspond with indexes in $\mathscr{A}_{(1)}$. The equality sign in the null hypothesis as well as the inequality sign in the alternative hypothesis are considered again element-wise.

One can easily see that the hypotheses test in (11) differs from the original test formulated in (7) but there is still some common sense in the later formulation presented in (11). Indeed, if the true function is increasing the final fit should correspond and thus by rejecting the null we actually confirm the imposed shape constraints. On the other hand, if the true function is rather decreasing, the final fit will result in a simple overall mean and under the imposed restrictions the null hypothesis should not be rejected as the true parameters are supposed to be zeros indeed. An analogous argumentation can be also used in case of some isotonic shape restrictions.

As far as we assume only limited, a small number of slope breaks to occur in the final model we have that $|\mathscr{A}| = \#\{j \in \mathscr{J};\ \hat{\beta}_j \neq 0\}$ is small and thus we only have a small number of the element-wise inequalities in (11). Using the post-selection approach discussed in Sect. 3.2 we only need to consider each inequality in a separate test and we can calculate the corresponding p-value for each test separately: applying Theorem 1 we can use the test statistic defined in (10) to calculate the corresponding p-values for individual tests $H_{0j} : \boldsymbol{a}_j^\top \boldsymbol{\beta}_{\mathscr{A}_{(1)}} = 0$ against $H_1 : \boldsymbol{a}_j^\top \boldsymbol{\beta}_{\mathscr{A}_{(1)}} > 0$, for each $j = 1, \ldots, J$, where $\boldsymbol{a}_j^\top$ are the corresponding columns of $\boldsymbol{A}_{\mathscr{A}_{(1)}} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p)^\top$ and $J = |\mathscr{A}_{(1)}|$.

Note, however, that the overall slope parameter β_1 is not penalized in the LASSO penalty in (3) and thus, there is no need to apply the polyhedral conditioning for calculating the corresponding p-value for this parameter. Indeed, it is easy to see from the KKT conditions that the estimate of the overall slope parameter is given in a sense of an ordinary linear regression and therefore, classical inference techniques can be applied. For other parameters which are penalized in the LASSO penalty we need

[7] In case of the isotonic constraints where the intercept parameter does not play any role in the test one can consider the set of indexes for active parameters only, the set $\mathscr{A}$.

to adjust for the randomness in the model selection step and thus, the corresponding p-values need to be obtained using the post-selection inference approach.

To conclude, we need to deal with a relatively small set of multiple tests while controlling for the overall first type error probability. Multiple comparisons techniques need to be used but one needs to be aware of the fact that given the structure of the matrix in (11) these multiple tests are not independent. There are many approaches and different techniques which can be applied to correctly control and manage the overall first type error probability: the simplest approach is so called Bonferroni correction which is however, known to be quite conservative in many situations. Another techniques are discussed, e.g. in [17] and more sophisticated approaches based on nonparametric combination of tests are summarized in [18]. For our purposes we use a simple modification of the Bonferroni correction, so called Holm-Bonferroni correction (see [9] for details): for $J \in \mathbb{N}$ being the number of comparisons in (11) (equivalently the number of rows in $\boldsymbol{A}_{\mathscr{A}_{(1)}}$) and $\alpha \in (0, 1)$ being the overall first type error probability we sort the corresponding p-values $p_{(1)} \leq \cdots \leq p_{(J)}$ and we search for the smallest $j \in \{1, \ldots, J\}$ such that $p_{(j)} > \alpha/(J - j + 1)$. If $j = 1$ then we do not reject the null hypothesis. If there is no such $j \in \{1, \ldots, J\}$ then we reject the null hypothesis in (11) and thus, we confirm the imposed shape constraints.

4 Finite Sample Results

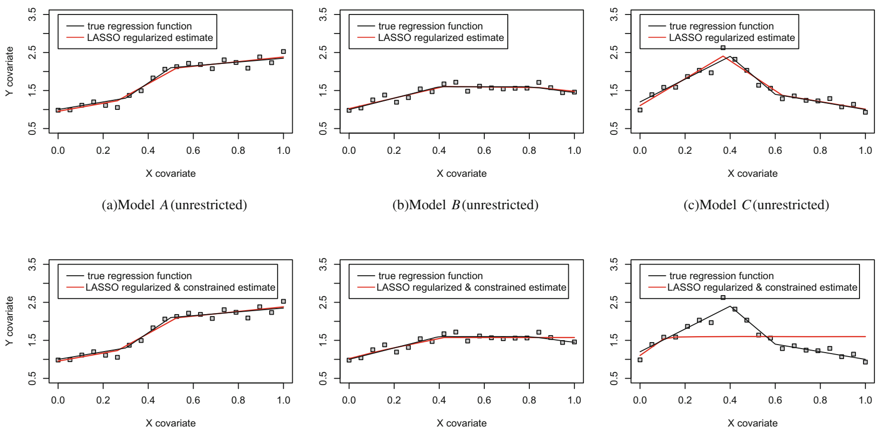
In this section we investigate some final sample properties of the proposed estimation and testing approach. Specifically, we are interested in estimating some unknown regression function which is assumed to be piecewise linear and continuous on its domain and in addition we impose non-decreasing[8] restriction on the constructed fit. Using the estimated model we would like use the proposed test to verify whether the shape restrictions are likely to be fulfilled or not.

For this purpose we consider three different models with piecewise linear and continuous underlying regression functions:

$$\begin{aligned}
\text{Model } A: &\quad f_1(x) = 1 + x\mathbb{I}_{(x<0.3)} + (4x - 0.9)\mathbb{I}_{(x\in[0.3,0.5))} + (0.5x + 0.85)\mathbb{I}_{(x\geq 0.5)};\\
\text{Model } B: &\quad f_2(x) = 1 + 1.5x\mathbb{I}_{(x<0.4)} + 0.6x\mathbb{I}_{(x\in[0.4,0.8))} - (0.75x - 3.2)\mathbb{I}_{(x\geq 0.8)};\\
\text{Model } C: &\quad f_3(x) = (3x + 1.2)\mathbb{I}_{(x<0.4)} - (5x - 4.4)\mathbb{I}_{(x\in[0.4,0.8))} - (x - 2)\mathbb{I}_{(x\geq 0.8)};
\end{aligned}$$

It is easy to verify that Model A satisfies the non-decreasing property over the whole domain— interval $(0, 1)$; Model B is increasing in the first part of the domain, constant in the second part and finally, slightly decreasing in the last part of the domain; Finally, Model C is similar as the second one but the decreasing segments are much more obvious, therefore, there should be more evidence in favour of the null

[8]One can assume different shape constraints. In this paper, however, we only present a small part of the simulation results for this specific restriction.



(a)Model A(unrestricted) (b)Model B(unrestricted) (c)Model C(unrestricted)

(d)Model A(shapeconstrained) (e)Model B(shapeconstrained) (f)Model C(shapeconstrained)

Fig. 1 Three underlying models (true functions with *black solid lines*) fitted with the LASSO regularized joinpoint regression approach (*red solid lines*): firstly, under no shape restrictions given by (3) and displayed in subfigures **a**–**c**; secondly, under the additional non-decreasing shape restriction as defined by (6) given in subfigures **d**–**f**

hypothesis (alternative means non-decreasing). The data were generated using the corresponding model where $n = 20$ design points were equidistantly placed over the whole domain and the random noise followed the normal distribution $N(0, 0.1^2)$. All design points were assumed to be hypothetical change-point locations for the slope break, however, we apriori considered two change-points only (resp. slope breaks) and we used the unrestricted estimation approach as defined in (3) (see Fig. 1a–c) and the shape-constrained estimation procedure in sense of (6) (Fig. 1d–f). Both estimation approaches are used to compare the effect of imposing the additional shape constraints.

It is clear from Fig. 1 that once the constraints are true, the model estimated under the constraints and the model estimated without constraints are identical (compare Fig. 1a, d). On the other hand, if the imposed constraints are validated by the true model then these two estimation approaches clearly differ. Indeed, one model tries to fit the underlying data and to reveal the true structure while the second one stays limited within the imposed restrictions (compare Fig. 1b, e and c, f).

As we are also interested in some statistical decision whether given constraints hold or not we consider a set of hypothesis defined in (11). We assume only two slope breaks by default, therefore, there are only three rows in $\boldsymbol{A}_{\mathscr{A}_{(1)}}$ and three non-zero parameters $\widehat{\boldsymbol{\beta}}_{\mathscr{A}_{(1)}} = (\widehat{\beta}_1 \widehat{\beta}_{j_1}, \widehat{\beta}_{j_2})^\top$ (the overall intercept and two slope breaks). Next, we calculate the corresponding p-values for simple tests of linear combinations of parameters with each line of the matrix $\boldsymbol{A}_{\mathscr{A}_{(1)}} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$.

Thus we obtain three p-values one for each row in $\boldsymbol{A}_{\mathscr{A}_{(1)}}$: the first corresponds with the partial test $H_{01} : \beta_1 = 0$ against $H_{11} : \beta_1 > 0$ for the overall slope parameter β_1

Table 1 Summary table for the partial statistical tests defined by the set of hypotheses in (11) given for all three considered models. Simulation results given for 100 independent repetitions where each model was re-fitted 100 times given 100 simulated datasets are also provided

Parameter estimates	Model *A*			Model *B*			Model *C*		
	p-value	Avg. (per 100)	Std.Err. (per 100)	p-value	Avg. (per 100)	Std.Err. (per 100)	p-value	Avg. (per 100)	Std.Err. (per 100)
$\widehat{\beta}_1$	**0.0085**	0.0111	(0.0029)	**0.1555**	0.2584	(0.0356)	**0.8253**	0.8447	(0.0255)
$\widehat{\beta}_1 + \widehat{\beta}_{j_1}$	**<0.0001**	0.0004	(0.0001)	**0.2830**	0.2569	(0.1244)	**0.6223**	0.5982	(0.2219)
$\widehat{\beta}_1 + \widehat{\beta}_{j_1} + \widehat{\beta}_{j_2}$	**0.0001**	0.0001	(0.0004)	**0.3321**	0.3848	(0.1829)	**0.5424**	0.5567	(0.3220)

and the remaining two correspond with analogous tests for the linear combinations $\beta_1 + \beta_{j_1}$ and $\beta_1 + \beta_{j_1} + \beta_{j_2}$, where $j_1, j_2 \in \mathscr{A}$ are the only two indexes of two active parameters. The corresponding p-values are calculated via the post-selection framework using Theorem 1. The corresponding p-values for each of the three models (A, B and C) are presented in Table 1.

All three models were also considered repeatedly in additional simulations in order to investigate the overall behaviour: the results are averaged for 100 repetitions and they are given in the last two columns of the corresponding partial table (one for each model) in Table 1.

To finally conclude, for $\alpha = 0.05$ the null hypothesis in case of Model A is clearly rejected in favour of the alternative which states that $\beta_1 > 0$ as well as $\beta_1 + \beta_{j_1} > 0$ and $\beta_1 + \beta_{j_1} + \beta_{j_2 0} > 0$ and thus, all three slope parameters of the three segments in the model are positive. For the remaining two models the null hypothesis in (11) is not rejected. Thus, the estimated parameters which would otherwise cause the estimated function to take a decreasing shape can be considered to be zero. Comparing this with the shape-constrained fits in Fig. 1e, f we can conclude that the slope parameters are mostly indeed being estimated as zeros.

5 Conclusion

Piecewise continuous linear regression models are popular modelling approaches in various practical applications where the final interpretation of the model should be kept as easy as possible. We considered the LASSO regularized estimation framework for fitting such models. However, due to an evident lack of sufficient statistical theory in the area of the LASSO estimation it is not easy to approach such models from the perspective of some valid inference.

In this paper we proposed an adaptation of the idea being recently presented in [20] and we introduced a statistical test which can be used to verify some qualitative properties of the final fit. Specifically, we consider some common shape constraints like monotonic or isotonic properties and we use the proposed statistical testing approach to validate such assumptions.

There are still some limitations involved in both, theoretical and practical aspects of this methodology. First of all, we can not directly consider the statistical test defined by hypotheses in (7) as we can only deal with simple null hypotheses as formulated for instance in (8). On the other hand, however, we showed that there is still some common sense even behind the later formulation and the problem can be still adopted to verify the imposed structural constraints in the joinpoint model: one just need to reconsider the problem and instead of testing the validity of the shape constraints in the null hypothesis as stated in (7) we rather test the null hypothesis on the corresponding parameters begin zero (meaning that the final model fit under the null will coincide with the considered shape restriction even though the true model may not) against the alternative that the model fit truly estimates the unknown model and they both comply with the given shape restriction.

There is still a way to improve the proposed testing approach in order to provide a more complex approach similar to, e.g. an F test or likelihood ratio test which are both effectively used in the classical linear regression framework. We believe however, that the proposed multi-stage test discussed in this paper can be still used in some practical situations where a decision on some imposed constraints needs to be done.

Acknowledgements We thank Ivan Mizera, Sen Bodhisattva and the referee for their comments and remarks. This work was partially supported by the PRVOUK grant 300-04/130.

References

1. Bosetti, C., Bertuccio, P., Levi, F., Lucchini, F., Negri, E., La Vecchia, C.: Cancer mortality in the European Union, 1970–2003, with a joinpoint analysis. Ann. Oncol. **16**, 631–640 (2008)
2. Carlin, B.P., Gelfand, A.E., Smith, A.F.M.: Hierarchical Bayesian analysis of changepoint problems. Appl. Stat. **41**, 389–405 (1992)
3. Efron, B., Hastie, T., Tibshirani, R.: Least angle regression. Ann. Stat. **32**, 407–499 (2004)
4. Feder, P.: The log likelihood ratio in segmented regression. Ann. Stat. **3**, 84–97 (1975)
5. Friedman, J., Hastie, T., Tibshirani, R.: Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. **33**, 1–22 (2010)
6. Gareth, M.J., Paulson, C., Rusmevichientong, P.: The Constrained Lasso. (working paper)
7. Harchaoui, Z., Lévy-Leduc, C.: Multiple change-point estimation with a total variation penalty. J. Am. Stat. Assoc. **105**(492), 1480–1493 (2010)
8. Hinkley, D.V.: Inference in two-phase regression. J. Am. Stat. Assoc. **66**(336), 736–743 (1971)
9. Holm, S.: A simple sequentially rejective multiple test procedure. Scand. J. Stat. **6**(2), 65–70 (1979)
10. Kim, H.J., Yu, B., Feuer, E.J.: Selecting the number of change-poins in segmented line regression. Stat. Sin. **19**, 597–609 (2009)
11. Kim, H.J., Fay, M.P., Feuer, E.J., Midthune, D.N.: Permutation tests for joinpoint regression with application to cancer rates. Stat. Med. **19**, 335–351 (2000)
12. Lee, J.D., Sun, D.,L., Sun, Y., Taylor, J.,E.: Exact post-selection inference with application to the lasso. Ann. Stat. **44**(3), 907–927 (2016)
13. Lockhart, R., Taylor, J., Tibshirani, R.J., Tibshirani, R.: A significance test for LASSO. Ann. Stat. **42**(2), 413–468 (2014)
14. Maciak, M., Mizera, I.: Regularization techniques in joinpoint regression. Statistical Papers (in revision) (2016)
15. Martinez-Beneito, M., García-Donato, G., Salmerón, D.: A bayesian joinpoint regression model with an unknown number of break-points. Ann. Appl. Stat. **5**(3), 2150–2168 (2011)
16. Meinshausen, N., Bühlmann, P.: Stability selection. J. R. Stat. Soc. B **72**(4), 417–473 (2010)
17. Miller, R.: Simultaneous Statistical Inference, 2nd edn. Springer, New York (1981)
18. Pesarin, F.: Multivariate Permutations Tests with Application in Biostatistics. Wiley, New York (2001)
19. Qui, D., Katanoda, K., Tomomi, M., Tomotaka, S.: A joinpoint regression analysis of long-term trends in cancer mortality in Japan (1958–2004). Int. J. Cancer **24**, 443–448 (2009)
20. Tibshirani, R.: Regression shrinkage and selection via the LASSO. J. R. Stat. Soc. Ser. B **58**(1), 267–288 (1996)

21. Tibshirani, R., Taylor, J., Lockhart, R., Tibshirani. R.: Exact post-selection inference for sequential regression procedures. J. Am. Stat. Assoc. (to appear)
22. van de Geer, S., Bühlman, P., Ritov, Y., Dezeure, R.: On asymptotically optimal confidence regions and tests for high-dimensional models. Ann. Stat. **42**(3), 1166–1201 (2014)
23. Zhang, C.H., Zhang, S.: Confidence intervals for low dimensional parameters in high dimensional linear models. J. R. Stat. Soc.: Ser. B **76**(1), 217–242 (2014)

Shape Constrained Regression in Sobolev Spaces with Application to Option Pricing

Michal Pešta and Zdeněk Hlávka

Abstract A class of nonparametric regression estimators based on penalized least squares over the sets of sufficiently smooth functions is elaborated. We impose additional shape constraint—isotonia—on the estimated regression curve and its derivatives. The problem of searching for the best fitting function in an infinite dimensional space is transformed into a finite dimensional optimization problem making this approach computationally feasible. The form and properties of the regression estimator in the Sobolev space are investigated. An application to option pricing is presented. The behavior of the estimator is improved by implementing an approximation of a covariance structure for the observed intraday option prices.

Keywords Constrained regression · Sobolev spaces · Isotonic constraints · Monotonicity · Covariance structure · Option price

1 Introduction

Suppose we need to estimate an unknown regression curve and the only restriction is that it should be *sufficiently smooth*. Therefore, we do not want to prescribe any functional form and a nonparametric approach is suitable. For obtaining a reasonable fit with plausible smoothness, *penalized least squares* are chosen, cf. [12]. We are interested in a regression model based on functions with specific features (quality).

An estimator of the regression curve will be an element of a *Sobolev space*, a class of functions with smooth high-order derivatives. Such approach was proposed for

M. Pešta (✉) · Z. Hlávka
Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Charles University, Sokolovská 49/83, 18675 Prague, Czech Republic
e-mail: Michal.Pesta@mff.cuni.cz

Z. Hlávka
e-mail: Zdenek.Hlavka@mff.cuni.cz

J. Antoch et al. (eds.), *Analytical Methods in Statistics*, Springer Proceedings in Mathematics & Statistics 193, DOI 10.1007/978-3-319-51313-3_7

random design by [13]. Nevertheless, considering the option pricing problem, the *fixed design* seems to be more appropriate.

Our primary goal is to investigate the form, asymptotic properties, and computational feasibility of the estimated regression curve in the Sobolev space. In addition, we also include *shape constraints* on the regression curve following from the put-call option price duality: obviously, isotonicity, i.e., [3], is nonnegativity (or nonpositivity) of certain higher order derivatives (e.g., monotonicity, convexity, or concavity).

The investigation is based on a shape-constrained regression in Sobolev spaces. In Sect. 2, we describe the mathematical foundation of the method. In Sect. 3, we introduce the estimator and investigate its consistency and computational availability. Specific problems arising in the real-life application on the observed option prices are discussed in Sect. 4 considering also an approximation to covariance structure for the observed call and put option prices proposed in Sect. 5. Finally, state price density (SPD) estimates based on the observed DAX option prices are calculated in Sect. 6. The proofs of all theorems are given in Appendix.

2 Sobolev Spaces' Framework

In this section, necessary preliminaries and some theorems for statistical regression in Sobolev spaces are assembled. The crux of this section lies in Theorem 1 (Representors in Sobolev Space). Compared to [13], we examine the so-called representors in more detail (see also Proof of Theorem 1).

The symbol $\mathsf{L_p}(\Omega)$ shall denote the *Lebesgue space* $\mathsf{L_p}(\Omega) := \{f : \|f\|_{\mathsf{L_p}(\Omega)} < \infty\}$, $1 \leq p \leq \infty$, where $\|f\|_{\mathsf{L_p}(\Omega)} := \left[\int_\Omega f^p(\mathbb{x})\mathrm{d}\mathbb{x}\right]^{1/p}$ for $1 \leq p < \infty$ and $\|f\|_{\mathsf{L_\infty}(\Omega)} := \inf\{C \geq 0 : |f| \leq C \text{ a.e.}\}$ for a measurable real-valued function $f : \Omega \to \mathbb{R}$ on a given Lebesgue-measurable domain Ω, i.e., a connected Lebesgue-measurable bounded subset of an Euclidean space $\mathbb{R}^q$ with nonempty interior. The symbol $\mathscr{C}^m(\Omega)$, $m \in \mathbb{N}_0$ denotes the *space of m-times continuously differentiable scalar functions* upon bounded domain Ω, i.e., $\mathscr{C}^m(\Omega) := \left\{f : \Omega \to \mathbb{R} \middle| D^{\boldsymbol{\alpha}} f \in \mathscr{C}^0(\Omega), |\boldsymbol{\alpha}|_\infty \leq m\right\}$, where $|\boldsymbol{\alpha}|_\infty = \max_{i=1,\ldots,q} |\alpha_i|$. Let us denote by $D^{\boldsymbol{\alpha}} f(\mathbb{x}) := \partial^{|\boldsymbol{\alpha}|_1} f(\mathbb{x})/\partial x_1^{\alpha_1} \ldots \partial x_q^{\alpha_q}$ the partial derivative of the function $f : \Omega \to \mathbb{R}$ in $\mathbb{x} \in \mathrm{int}(\Omega)(\equiv \Omega^\circ := \overline{\Omega}\backslash\partial\Omega)$, where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_q)^\top \in \mathbb{N}_0^q$ is a multi-index of the modulus $|\boldsymbol{\alpha}|_1 = \sum_{i=1}^q \alpha_i$.

Definition 1 (*Sobolev norm*) Let $f \in \mathscr{C}^m(\Omega) \cap \mathsf{L_p}(\Omega)$. We introduce a Sobolev norm $\|\cdot\|_{p,Sob,m}$ as

$$\|f\|_{p,Sob,m} := \left\{\sum_{|\boldsymbol{\alpha}|_\infty \leq m} \int_\Omega \left|D^{\boldsymbol{\alpha}} f(\mathbb{x})\right|^p \mathrm{d}\mathbb{x}\right\}^{1/p}, \quad \text{where } |\boldsymbol{\alpha}|_\infty = \max_{i=1,\ldots,q} \alpha_i. \quad (1)$$

The Sobolev norm is a correctly defined norm, because the triangle inequality for the Sobolev norm (1) follows easily from the triangle inequality for the p-norms on $\mathsf{L}_\mathsf{p}(\Omega)$ and $l_p\,(\{\boldsymbol{\alpha} : |\boldsymbol{\alpha}|_\infty \leq m\})$.

Definition 2 (*Sobolev space*) A Sobolev space of rank m, $\mathscr{W}_p^m(\Omega)$, is the completion of the intersection of $\mathscr{C}^m(\Omega)$ and $\mathsf{L}_\mathsf{p}(\Omega)$ with respect to the Sobolev norm $\|\cdot\|_{p,Sob,m}$.

Note that $\mathscr{C}^m(\Omega) \cap \mathsf{L}_\mathsf{p}(\Omega)$ is dense in $\mathscr{W}_p^m(\Omega)$ according to $\|\cdot\|_{p,Sob,m}$.

Definition 3 (*Sobolev inner product*) Let $f, g \in \mathscr{W}_2^m(\Omega)$. The Sobolev inner product $\langle\cdot,\cdot\rangle_{Sob,m}$ is defined as

$$\langle f, g\rangle_{Sob,m} := \sum_{|\boldsymbol{\alpha}|_\infty \leq m} \int_\Omega D^{\boldsymbol{\alpha}} f(\mathbb{x}) D^{\boldsymbol{\alpha}} g(\mathbb{x}) \mathrm{d}\mathbb{x}.$$

The correctness of Definition 3 is guaranteed by the denseness of the space $\mathscr{C}^m(\Omega) \cap L_2(\Omega)$ in $\mathscr{W}_2^m(\Omega)$. The Sobolev inner product $\langle\cdot,\cdot\rangle_{Sob,m}$ induces the Sobolev norm $\|\cdot\|_{2,Sob,m}$ in $\mathscr{W}_2^m(\Omega)$ and we denote the Sobolev space $\mathscr{H}^m(\Omega) := \mathscr{W}_2^m(\Omega)$. For simplicity of notation, we denote the Sobolev norm $\|\cdot\|_{2,Sob,m} := \|\cdot\|_{Sob,m}$. It is straightforward to verify that $\mathscr{H}^m(\Omega)$ is a normed linear space. By construction, $\mathscr{H}^m(\Omega)$ is complete and, hence, it is a Banach space. Next, the inner product $\langle\cdot,\cdot\rangle_{Sob,m}$ has been defined on $\mathscr{H}^m(\Omega)$ and it follows that $\mathscr{H}^m(\Omega)$ is a Hilbert space.

2.1 Construction of Representors in Sobolev Space

The Hilbert space $\mathscr{H}^m(\Omega)$ can be expressed as a direct sum of subspaces that are orthogonal to each other. For the nonparametric regression, see Sect. 3, it is very important that we can take projections of the elements of $\mathscr{H}^m(\Omega)$ into its subspaces.

The following Theorem 1 is the representation theorem for Sobolev spaces derived in [13], an analogy to the well-known Riesz representation theorem. From now on, we suppose that $m \in \mathbb{N}$. The symbol $\mathscr{Q}^q$ denotes the closed unit cube in $\mathbb{R}^q$.

Theorem 1 (Representors in Sobolev space) *For all* $f \in \mathscr{H}^m(\mathscr{Q}^q)$, $\mathbb{a} \in \mathscr{Q}^q$ *and* $\mathbb{w} \in \mathbb{N}_0^q$, $|\mathbb{w}|_\infty \leq m-1$, *there exists a representor* $\psi\ \ (\mathbb{x}) \in \mathscr{H}^m(\mathscr{Q}^q)$ *at the point* $\mathbb{a}$ *with the rank* $\mathbb{w}$ *such that* $\langle \psi\ \ , f\rangle_{Sob,m} = D\ \ f(\mathbb{a})$. *Moreover,* $\psi\ \ (\mathbb{x}) = \prod_{i=1}^q \psi_{a_i}^{w_i}(x_i)$ *for all* $\mathbb{x} \in \mathscr{Q}^q$, *where* $\psi_{a_i}^{w_i}(\cdot)$ *is the representor in the Sobolev space of functions of one variable on* $\mathscr{Q}^1$ *with the inner product*

$$\begin{aligned}\frac{\partial^{w_i} f(\mathbb{a})}{\partial x_i^{w_i}} &= \left\langle \psi_{a_i}^{w_i}, f(x_1,\ldots,x_{i-1},\cdot,x_{i+1},\ldots,x_q)\right\rangle_{Sob,m} \\ &= \sum_{\alpha=0}^m \int_{\mathscr{Q}^1} \frac{d^\alpha \psi_{a_i}^{w_i}(x_i)}{dx_i^\alpha} \frac{d^\alpha f(\mathbb{x})}{dx_i^\alpha} dx_i. \qquad (2)\end{aligned}$$

The proof of Theorem 1 given in Appendix is based on the ideas of [13]. In addition, we derive the *exact form of the representor* for Sobolev spaces $\mathscr{H}^m(\Omega)$ in this proof, which provides the *computational feasibility* of the whole approach.

In order to derive the representor $\psi_a \equiv \psi_a^0$ of $f \in \mathscr{H}^m\,[0, 1]$, we start with functions L_a and R_a defined in Appendix in (47)–(50). The coefficients $\gamma_k(a)$ of the representor are obtained as the solution of a system of linear equations corresponding to the boundary conditions (35)–(39) of the differential equation (34). The existence and uniqueness of the coefficients $\gamma_k(a)$ are shown in the proof of Theorem 1.

Theorem 2 (Obtaining coefficients $\gamma_k(a)$) *The coefficients $\gamma_k(a)$ of the representor ψ_a are the unique solution of the following $4m \times 4m$ system of linear equations*

$$\sum_{\substack{k=0\\k\neq\kappa}}^{m} \gamma_k(a)\left\{\varphi_k^{(m-j)}(0)+(-1)^j\varphi_k^{(m+j)}(0)\right\}$$
$$+\sum_{\substack{k=0\\k\neq\kappa}}^{m} \gamma_{m+1+k}(a)\left\{\varphi_{m+1+k}^{(m-j)}(0)+(-1)^j\varphi_{m+1+k}^{(m+j)}(0)\right\}=0 \quad \textit{for} \quad j=0,\ldots,m-1, \tag{3}$$

$$\sum_{\substack{k=0\\k\neq\kappa}}^{m} \gamma_{2m+2+k}(a)\left\{\varphi_k^{(m-j)}(1)+(-1)^j\varphi_k^{(m+j)}(1)\right\}$$
$$+\sum_{\substack{k=0\\k\neq\kappa}}^{m} \gamma_{3m+3+k}(a)\left\{\varphi_{m+1+k}^{(m-j)}(1)+(-1)^j\varphi_{m+1+k}^{(m+j)}(1)\right\}=0 \quad \textit{for} \quad j=0,\ldots,m-1, \tag{4}$$

$$\sum_{\substack{k=0\\k\neq\kappa}}^{m} \{\gamma_k(a)-\gamma_{2m+2+k}(a)\}\,\varphi_k^{(j)}(a)$$
$$+\sum_{\substack{k=0\\k\neq\kappa}}^{m} \{\gamma_{m+1+k}(a)-\gamma_{3m+3+k}(a)\}\,\varphi_{m+1+k}^{(j)}(a)=0, \quad \textit{for} \quad j=0,\ldots,2m-2, \quad \textit{and} \tag{5}$$

$$\sum_{\substack{k=0\\k\neq\kappa}}^{m} \{\gamma_k(a)-\gamma_{2m+2+k}(a)\}\,\varphi_k^{(2m-1)}(a)$$
$$+\sum_{\substack{k=0\\k\neq\kappa}}^{m} \{\gamma_{m+1+k}(a)-\gamma_{3m+3+k}(a)\}\,\varphi_{m+1+k}^{(2m-1)}(a)=(-1)^{m-1}, \tag{6}$$

where κ is the integer part of $(m+1)/2$ and φ_k are defined in Appendix in (41)–(46).

Let 0_n denote column vector of zeros of length n and $\boldsymbol{\gamma}(a)$ a column vector of the coefficients $\gamma_k(a)$ from Eqs. (3)–(6), i.e.,

$$\boldsymbol{\gamma}(a) = \big[\gamma_0(a), \dots, \gamma_{\kappa-1}(a), \gamma_{\kappa+1}(a), \dots, \gamma_{m+\kappa}(a), \gamma_{m+2+\kappa}(a), \dots, \gamma_{2m+1+\kappa}(a), \\ \gamma_{2m+3+\kappa}(a), \dots, \gamma_{3m+2+\kappa}(a), \gamma_{3m+4+\kappa}(a), \dots, \gamma_{4m+3}(a)\big]^\top.$$

The system of the $4m$ linear equations (3)–(6) can now be written in a more illustrative way

$$\underbrace{\left[\begin{array}{c|c} \varphi_k^{(m-j)}(0) + (-1)^j \varphi_k^{(m+j)}(0) & 0_{m-1}0_{m-1}^\top \\ \hline 0_{m-1}0_{m-1}^\top & \varphi_k^{(m-j)}(1) + (-1)^j \varphi_k^{(m+j)}(1) \\ \hline \varphi_k^{(j)}(a) & -\varphi_k^{(j)}(a) \\ \hline \varphi_k^{(2m-1)}(a) & -\varphi_k^{(2m-1)}(a) \end{array}\right]}_{\{\Gamma_{j,k}(a)\}_{j,k}} \boldsymbol{\gamma}(\alpha) = \left[\begin{array}{c} 0_m \\ \hline 0_m \\ \hline 0_{2m-1} \\ \hline (-1)^{m-1} \end{array}\right]. \tag{7}$$

Hence, the coefficients can be expressed as $\boldsymbol{\gamma}(a) = (-1)^{m-1}\left[\{\boldsymbol{\Gamma}(a)\}^{-1}\right]_{\bullet,4m}$, which provides a direct way for the calculation of the representors. Furthermore, the form of representor ψ $(\mathbb{x})$ does not depend on the boundaries of the closed cube. In Fig. 1, one can see representors in Sobolev space for 11 data points.

3 Penalized Least Squares

A combination of properties of the L_2 and $\mathscr{C}^m$ space yields an interesting background for nonparametric regression. The L_2 space is a special type of Hilbert space that facilitates the calculation of least squares projection. On the other hand, the $\mathscr{C}^m$ space contains classes of smooth (m-times continuously differentiable) functions suitable for nonparametric regression.

The investigated *regression model* is

$$Y_i = f(\mathbb{x}_i) + \varepsilon_i, \quad i = 1, \dots, n, \tag{8}$$

where $\mathbb{x}_i$ are q-dimensional fixed design points (knots), ε_i are correlated random errors such that $\mathsf{E}\varepsilon_i = 0$ and $\mathsf{Var}\varepsilon = \boldsymbol{\Sigma} = (\sigma_{ij})_{i,j=1,\dots,n}$ with $\sigma_i^2 = \sigma_{ii} > 0$, and $f \in \mathscr{F}$, where $\mathscr{F}$ is a family of functions in the Sobolev space $\mathscr{H}^m(\mathscr{Q}^q)$ from $\mathbb{R}^q$ to $\mathbb{R}^1$, $m > \frac{q}{2}$, $\mathscr{F} = \left\{f \in \mathscr{H}^m(\mathscr{Q}^q) : \|f\|_{Sob,m}^2 \le L\right\}$. From now on, we denote $\mathscr{H}^m \equiv \mathscr{H}^m(\mathscr{Q}^q)$.

For heteroscedastic and correlated data, the estimation of f is carried out as

$$\widehat{f} = \arg\min_{f \in \mathscr{H}^m} \frac{1}{n}\left[\mathbb{Y} - \boldsymbol{f}(\mathbb{x})\right]^\top \boldsymbol{\Sigma}^{-1}\left[\mathbb{Y} - \boldsymbol{f}(\mathbb{x})\right] + \chi\,\|f\|_{Sob,m}^2, \tag{9}$$

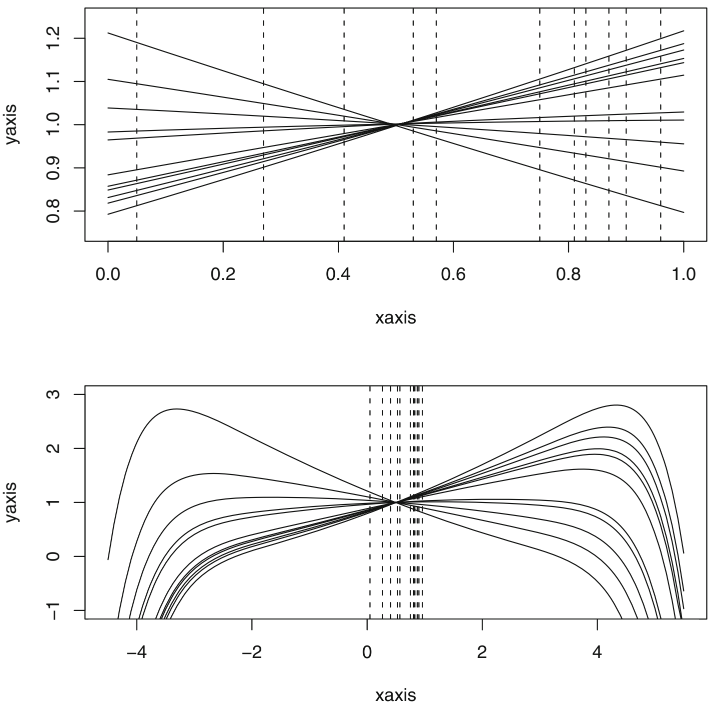


Fig. 1 The *full lines* in both plots are the representors in the Sobolev space $\mathscr{H}^4[0,1]$ for data points $\mathbb{x} = [0.05, 0.27, 0.41, 0.53, 0.57, 0.75, 0.81, 0.83, 0.87, 0.9, 0.96]^\top$ marked by *dashed vertical lines*. The limits of the horizontal axis in the upper and the lower graph are $[0, 1]$ and $[-4.5, +5.5]$, respectively

where $\mathbb{x}$ is $(n \times q)$ matrix containing in its rows the q-dimensional design points $\mathbb{x}_1, \ldots, \mathbb{x}_n$, $\boldsymbol{\Sigma} > 0$ is $n \times n$ symmetric (variance) matrix, $\mathbb{Y}$ is $n \times 1$ vector of observations, $\boldsymbol{f}(\mathbb{x}) = [f(\mathbb{x}_1), \ldots, f(\mathbb{x}_n)]^\top$, and $\chi > 0$.

In order to derive the estimator of the unknown regression curve, one needs to define a representor matrix.

Definition 4 (*Representor matrix*) Let $\psi_1, \ldots, \psi_n$ be the representors for function evaluation at $\mathbb{x}_1, \ldots, \mathbb{x}_n$, respectively. i.e., $\left\langle \psi_i, f \right\rangle_{Sob,m} = f(\mathbb{x}_i)$ for all $f \in \mathscr{H}^m$, $i = 1, \ldots, n$. The representor matrix $\boldsymbol{\Psi}$ is the $(n \times n)$ matrix such that its columns and rows are the representors evaluated at $\mathbb{x}_1, \ldots, \mathbb{x}_n$, i.e., $\boldsymbol{\Psi} = (\Psi_{i,j})_{i,j=1,\ldots,n}$, where $\Psi_{i,j} = \left\langle \psi_i, \psi_j \right\rangle_{Sob,m} = \psi_i(\mathbb{x}_j) = \psi_j(\mathbb{x}_i)$.

The forthcoming theorem transforms the *infinite dimensional problem into a finite dimensional quadratic optimization problem*, which makes the approach computationally efficient. Similar result derived by [13] uses different penalization.

Theorem 3 (Infinite to finite) *Assume that* $\mathbb{Y} = [Y_1, \ldots, Y_n]^\top$ *and* $\boldsymbol{\Sigma} > 0$ *is* $(n \times n)$ *symmetric matrix. Define*

$$\hat{\sigma}^2 = \min_{f \in \mathscr{H}^m} \frac{1}{n} \left[\mathbb{Y} - \boldsymbol{f}(\mathbb{x})\right]^\top \boldsymbol{\Sigma}^{-1} \left[\mathbb{Y} - \boldsymbol{f}(\mathbb{x})\right] + \chi \, \|f\|^2_{Sob,m}\,, \tag{10}$$

$$s^2 = \min_{\in \mathbb{R}^n} \frac{1}{n} \left[\mathbb{Y} - \boldsymbol{\Psi}\mathbb{c}\right]^\top \boldsymbol{\Sigma}^{-1} \left[\mathbb{Y} - \boldsymbol{\Psi}\mathbb{c}\right] + \chi \mathbb{c}^\top \boldsymbol{\Psi} \mathbb{c} \tag{11}$$

where $\mathbb{c}$ *is a* $(n \times 1)$ *vector,* $\boldsymbol{f}$ *is defined in* (9)*, and* $\boldsymbol{\Psi}$ *is the representor matrix. Then* $\hat{\sigma}^2 = s^2$*. Furthermore, there exists a solution to* (10) *of the form* $\widehat{f} = \sum_{i=1}^n \hat{c}_i \psi_{\ \ i}$*, where* $\widehat{\mathbb{c}} = [\hat{c}_1, \ldots, \hat{c}_n]^\top$ *solves* (11)*. The estimator* $\widehat{f}$ *is unique a.e.*

Additionally, we give a *closed form* of the regression function estimator using the objects defined in Appendix.

Corollary 1 (Form of the regression function estimator) *In one-dimensional case* $(q = 1)$*, the regression function estimator* $\widehat{f}$ *defined in Theorem 3 can be written as*

$$\widehat{f}(x) = \begin{cases} \sum_{i=1}^{n} \hat{c}_i L_{x_i}(x), & 0 \le x \le x_1, \\ \sum_{i=j+1}^{n} \hat{c}_i L_{x_i}(x) + \sum_{i=1}^{j} \hat{c}_i R_{x_i}(x), & x_j < x \le x_{j+1}, \ j = 1, \ldots, n-1; \\ \sum_{i=1}^{n} \hat{c}_i R_{x_i}(x), & x_n < x \le 1, \end{cases} \tag{12}$$

where $\widehat{\mathbb{c}} = [\hat{c}_1, \ldots, \hat{c}_n]^\top$ *solves* (11) *and* $L_{x_i}(x)$ *and* $R_{x_i}(x)$ *are defined in* (30).

Corollary 1 can be easily extended for a q-dimensional vector variable $\mathbb{x}$ if we recall how the representor ψ is produced in the proof of Theorem 1. We apply (30) on the form of each factor ψ_a of the product of representors ψ as defined in part (ii) of the proof of Theorem 1. The only difference in (12) will be the number of cases. We will obtain $(n+1)^q$ decision conditions (vector $\mathbb{x}$ has q components) instead of actual number $n+1$ $(0 \le x \le x_1, \ldots, x_j < x \le x_{j+1}, \ldots, x_n < x)$.

Alternatively, the regression function estimator $\widehat{f}$ can be written as

$$\widehat{f}(x) = \sum_{j=1}^{n} \widehat{c_j} \sum_{k=1}^{2m} \exp\left[\Re\left(e^{i\theta_k}\right)x\right] \left\{ I_{[x \le x_j]} \gamma_k(x_j) \cos\left[\Im\left(e^{i\theta_k}\right)x\right] + I_{[x > x_j]} \gamma_{2m+k}(x_j) \sin\left[\Im\left(e^{i\theta_k}\right)x\right] \right\}.$$

Note that the estimator $\widehat{f}$ is not calculated using trigonometric splines neither kernel functions.

Theorem 4 (Symmetry and positive definiteness of representor matrix) *The representor matrix is symmetric and positive definite.*

In the linear model, the unknown coefficients are estimated using least squares. Gauss–Markov Theorem [10, Chap. 4] says that the least squares estimator is the best linear unbiased estimator and underlies the normal equations. The normal equations for our model are derived in Theorem 5.

Theorem 5 (Normal equations for $\widehat{\mathbb{c}}$) *Let us consider the general single equation model* (8). *Let* $\mathbb{Y}$ *denote the response vector* $[Y_1, \ldots, Y_n]^\top$ *and* $\boldsymbol{\Psi}$ *the representor matrix. Then, the vector* $\widehat{\mathbb{c}} = [c_1, \ldots, c_n]^\top$ *of the coefficients of the minimizer* $\widehat{f} = \sum_{i=1}^{n} \hat{c}_i \psi_i$ *derived in Theorem 3 is the unique solution of the system of equations* $\left(\boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi} + n\chi\boldsymbol{\Psi}\right)\mathbb{c} = \boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1}\mathbb{Y}$.

The fitted values $\widehat{\mathbb{Y}}$ can be expressed as $\widehat{\mathbb{Y}} = \widehat{\boldsymbol{f}}(\mathbb{x}) = \boldsymbol{\Psi}\widehat{\mathbb{c}}$. From the normal equations for $\widehat{\mathbb{c}}$ (Theorem 5), *hat matrix* $\boldsymbol{\Lambda} := \boldsymbol{\Psi}\left(\boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi} + n\chi\boldsymbol{\Psi}\right)^{-1}\boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1}$ satisfying $\widehat{\mathbb{Y}} = \boldsymbol{\Lambda}\mathbb{Y}$ can be obtained.

Using Theorem 3 and the Lagrange multipliers, a *one-to-one correspondence* between the *Sobolev bound* L and the *smoothing parameter* χ can be easily shown. Optimization problem (9) can be equivalently rewritten as

$$\arg\min_{f \in \mathscr{H}^m} \frac{1}{n}\left[\mathbb{Y} - \boldsymbol{f}(\mathbb{x})\right]^\top \boldsymbol{\Sigma}^{-1}\left[\mathbb{Y} - \boldsymbol{f}(\mathbb{x})\right] \quad \text{s.t.} \quad \|f\|^2_{Sob,m} \le L. \tag{13}$$

The Sobolev norm bound L and the smoothing (or bandwidth) parameter χ control the trade-off between the infidelity to the data and the roughness of the estimator.

The consistency result for the estimator is provided as a guarantee of the estimator's suitability.

Theorem 6 (Asymptotic behavior) *Suppose that* $\tilde{\boldsymbol{\varepsilon}} := \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\varepsilon}$ *is an* $(n \times 1)$ *vector of random variables. Then*

$$\frac{1}{n}\left[\widehat{\boldsymbol{f}}(\mathbb{x}) - \boldsymbol{f}(\mathbb{x})\right]^\top \boldsymbol{\Sigma}^{-1}\left[\widehat{\boldsymbol{f}}(\mathbb{x}) - \boldsymbol{f}(\mathbb{x})\right] = \mathscr{O}_{\mathsf{P}}\left(n^{-\frac{2m}{2m+q}}\right), \quad n \to \infty.$$

3.1 Choice of the Smoothing Parameter

The smoothing parameter χ corresponds to the diameter of the set of functions over which the estimation takes place. Heuristically, for large bounds (i.e., smaller χ), we obtain consistent but less efficient estimator. On the other hand, for smaller bounds (i.e., large χ) we obtain more efficient but inconsistent estimators.

A well-known selection method for the smoothing parameter χ is based on the minimization of the *cross-validation* criterion

$$\mathscr{CV}(\chi) = \frac{1}{n}\left[\mathbb{y} - \widehat{\boldsymbol{f}^*}(\mathbb{x})\right]^\top \boldsymbol{\Sigma}^{-1}\left[\mathbb{y} - \widehat{\boldsymbol{f}^*}(\mathbb{x})\right],$$

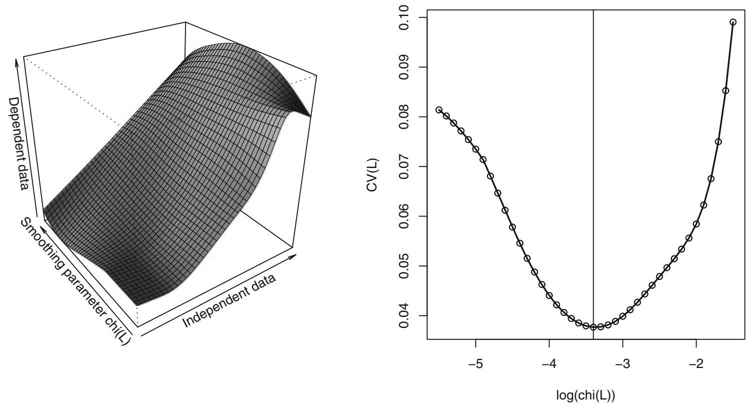


Fig. 2 The *left plot* shows how the fitted curve in $\mathscr{H}^2$ changes depending on the smoothing parameter χ. The *right plot* displays the cross-validation criterion as function of χ. The optimal value of the smoothing parameter is marked by a *vertical line*

where $\widehat{\boldsymbol{f}^*} = \left[\widehat{f}_{-1}, \ldots, \widehat{f}_{-n}\right]^\top$ is the usual leave-one-out estimator obtained by solving

$$\widehat{f}_{-i} = \arg \min_{f \in \mathscr{H}^m} \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^{n} \left[\boldsymbol{\Xi}_{j,\bullet}\mathbb{y} - \boldsymbol{\Xi}_{j,\bullet}\boldsymbol{f}(\mathbb{x})\right]^2 + \chi \, \|f\|^2_{Sob,m}, \quad i = 1, \ldots, n$$

and $\boldsymbol{\Xi}$ denotes the square root matrix of $\boldsymbol{\Sigma}^{-1}$. The smoothing parameter $\chi \equiv \chi(L)$, which in-turn corresponds to unique Sobolev bound L, is chosen as the minimizer of the cross-validation function $\mathscr{CV}$. Hence, the smoothing parameter χ (and the Sobolev bound L as well) is chosen based on the data prior to the estimation of f. The relationship between the fit and the smoothness of the estimator is plotted in Fig. 2.

Detailed information concerning the choice of the smoothing parameter χ can be found in [5]. Apart of the cross-validation, there exist many other methods based on penalizing functions or plug-in selectors.

4 Application to Option Prices

In Sect. 3, we have imposed only smoothness constraint on the estimated regression function $f \in \mathscr{F} = \left\{f \in \mathscr{H}^m(\mathscr{Q}^q) : \|f\|^2_{Sob,m} \leq L\right\}$. However, in practice we often have a prior knowledge concerning the shape of the regression function. In this section, we focus on the inclusion of additional constraints, such as isotonia or convexity, in the nonparametric regression estimator.

More formally, we are interested in the estimation of $f \in \tilde{\tilde{\mathscr{F}}} \subseteq \mathscr{F}$ where $\tilde{\tilde{\mathscr{F}}}$ combines smoothness with further properties such as monotonicity of particular derivatives of the function. The following discussion concerns only the one-dimensional case ($q = 1$).

Definition 5 (*Derivative of the representor matrix*) Let $\psi_{x_1}, \ldots, \psi_{x_n}$ be the representors for function evaluation at $x_1, \ldots, x_n$, i.e., $\langle \psi_{x_i}, f \rangle_{Sob,m} = f(x_i)$ for all $f \in \mathscr{H}^m$, $i = 1, \ldots, n$. The k-th derivative of the representor matrix $\boldsymbol{\Psi}$ is the matrix $\boldsymbol{\Psi}^{(k)}$ whose columns are equal to the k-th derivatives of the representors evaluated at $x_1, \ldots, x_n$, i.e., $\Psi^{(k)}_{i,j} = \psi^{(k)}_{x_j}(x_i)$, $i, j = 1, \ldots, n$.

Contrary to Theorem 4, the derivatives of the representor matrix do not have to be symmetric.

Definition 6 (*Estimate of the derivative*) The estimate of the derivative of the regression function is defined as the derivative of the regression function estimate, i.e., $\widehat{f^{(s)}} := \widehat{f}^{(s)}$, $s \in \mathbb{N}$.

Extending Theorem 6, the *uniform consistency* for the regression curve estimator and regression curve's derivatives estimator is proved for the fixed design setup.

Theorem 7 (Consistency of the estimator) *Suppose that* $\tilde{\boldsymbol{\varepsilon}} := \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\varepsilon}$ *is a* $(n \times 1)$ *vector of iid random variables, the design points are equidistantly distributed on the interval* $[a, b]$ *such that* $a = x_1 < \cdots < x_n = b$ *and* $\boldsymbol{\Sigma} > 0$ *is a covariance matrix of* $\boldsymbol{\varepsilon}$ *such that its largest eigenvalue is less or equal than a positive constant* $\vartheta > 0$ *for all* $n \in \mathbb{N}$. *Then* $\sup_{x \in [a,b]} \left| \widehat{f^{(s)}}(x) - f^{(s)}(x) \right| \xrightarrow[n\to\infty]{\mathsf{P}} 0$ *for* $s = 0, \ldots, m - 2$.

4.1 State Price Density

Let $Y_t(\varpi, T)$ denote the price of a European Call with strike price ϖ on day t and with expiry date T. The payoff at time T is given by $(S_T - \varpi)_+ = \max(S_T - \varpi, 0)$, where S_T denotes the price of the underlying asset at time T. The price of such an option may be expressed as the expected value of the payoff

$$Y_t(\varpi, T) = \exp\{-r(T - t)\} \int_0^{+\infty} (S_T - \varpi)_+ h(S_T) dS_T, \tag{14}$$

discounted by the known risk-free interest rate r. The expectation in (14) is evaluated with respect to the so-called *State Price Density (SPD)* $h(\cdot)$. The SPD contains important information on the expectations of the market and its estimation is a statistical task of great practical interest, see [8]. Similarly, we can express the price $Z_t(\varpi, T)$ of the European Put with payoff $(\varpi - S_T)_+$ as

$$Z_t(\varpi, T) = \exp\{-r(T-t)\} \int_0^{+\infty} (\varpi - S_T)_+ h(S_T) dS_T. \quad (15)$$

Calculating the second derivative of (14) and (15) with respect to the strike price ϖ, we can express the SPD as the second derivative of the European Call and Put option prices as in [4]:

$$h(\varpi) = \exp\{r(T-t)\} \frac{\partial^2 Y_t(\varpi, T)}{\partial \varpi^2} = \exp\{r(T-t)\} \frac{\partial^2 Z_t(\varpi, T)}{\partial \varpi^2}. \quad (16)$$

Both parametric and nonparametric approaches to the SPD estimation are described in [8]. Nonparametric estimates of the SPD based on (16) are considered, among others, in [1, 2, 6, 14]. Furthermore, the SPD was also estimated using the Kalman filter by [7], but the resulting estimate does not have to be smooth.

4.2 Call and Put Options

Suppose that Call and Put option prices are observed repeatedly for fixed distinct strike prices ϖ_i, $i = 1, \ldots, \omega$. The points ϖ_i are called the strike price knots.

In each strike price knot ϖ_i, we observe $n_i \in \mathbb{N}_0$ Call option prices Y_{i_k} with strikes $x_{i_k} = \varpi_i$, for $k = 1, \ldots, n_i$. We observe altogether $n = \sum_{i=1}^{\omega} n_i I_{[n_i \geq 1]}$ Call options in $\omega_Y = \sum_{i=1}^{\omega} I_{[n_i \geq 1]}$ distinct strike price knots. Similarly, in each strike price knot ϖ_j, $j = 1, \ldots, \omega$, we observe m_j Put option prices Z_{j_l} with strike prices $x_{j_l} = \varpi_j$, for $l = 1, \ldots, n_j$. In $\omega_Z = \sum_{j=1}^{\omega} I_{[m_j \geq 1]}$ distinct strike price knots, we observe $m = \sum_{j=1}^{\omega} m_j I_{[m_j \geq 1]}$ Put option prices.

Let us now denote by $\mathbb{Y}$ the vector of all observed Call option prices and by $\mathbb{x}_\alpha = (x_{\alpha,1}, \ldots, x_{\alpha,n})^\top$ the vector of the corresponding strike prices. Next, the symbol $\boldsymbol{\Delta} = (\Delta_{ij})_{i=1,\ldots,n;\, j=1\ldots,\omega_Y}$ denotes the connectivity matrix for Call option strike prices such that $\Delta_{ij} = I_{[x_{\alpha,i} = \varpi_j]}$. The symbol $\mathbb{Z}$ denotes the observed Put option prices. The vector $\mathbb{x}_\beta = (x_{\beta,1}, \ldots, x_{\beta,m})^\top$ of the strike prices corresponding to $\mathbb{Z}$ leads the connectivity matrix $\boldsymbol{\Theta} = (\Theta_{ij})_{i=1,\ldots,m;\, j=1\ldots,\omega_Z}$ for Put option prices defined as $\Theta_{ij} = I_{[x_{\beta,i} = \varpi_j]}$.

Our model for the observed Call and Put options prices can be written as:

$$Y_i = f(x_{\alpha,i}) + \varepsilon_i, \text{ where } x_{\alpha,i} \in \{\varpi_1, \ldots, \varpi_\omega\} \text{ and } i = 1, \ldots, n, \quad (17)$$

$$Z_j = g(x_{\beta,j}) + \nu_j, \text{ where } x_{\beta,j} \in \{\varpi_1, \ldots, \varpi_\omega\} \text{ and } j = 1, \ldots, m. \quad (18)$$

under the assumptions:

(i) ε_i and ν_j are random variables such that $\mathsf{E}\varepsilon_i = \mathsf{E}\nu_j = 0, \forall i, j$, $\mathsf{Cov}(\varepsilon_i, \varepsilon_k) = \xi_{i,k}$, $\mathsf{Cov}(\nu_j, \nu_l) = \zeta_{j,l}$, and $\mathsf{Cov}(\varepsilon_i, \nu_j) = \sigma_{i,j}$. For simplicity, we will write $\xi_i^2 = \xi_{i,i}$ and $\zeta_j^2 = \zeta_{j,j}$.

(ii) $f, g \in \mathscr{F}$, where $\mathscr{F} = \left\{ f \in \mathscr{H}^p : \|f\|^2_{Sob,p} \leq L \right\}$ and $p > \frac{1}{2}$.

We assume that the second derivatives of functions f and g have to be the same SPD, see equations (14)–(16) in the introduction. Theorem 3 allows to handle multiple (repeated) observations in our option prices setup (17)–(18).

Theorem 8 (Call and put option optimizing) *Invoke the assumptions from Call and Put Option Model* (17)–(18). *Define*

$$\hat{\sigma}^2 = \min_{f \in \mathscr{H}^p, g \in \mathscr{H}^p} \left[\begin{pmatrix} \mathbb{Y} \\ \mathbb{Z} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\Delta} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Theta} \end{pmatrix} \begin{pmatrix} \boldsymbol{f}\left(\mathbb{x}_\alpha\right) \\ \boldsymbol{g}\left(\mathbb{x}_\beta\right) \end{pmatrix} \right]^\top \boldsymbol{\Sigma}^{-1} \left[\begin{pmatrix} \mathbb{Y} \\ \mathbb{Z} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\Delta} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Theta} \end{pmatrix} \begin{pmatrix} \boldsymbol{f}\left(\mathbb{x}_\alpha\right) \\ \boldsymbol{g}\left(\mathbb{x}_\beta\right) \end{pmatrix} \right] + \chi \|f\|^2_{Sob,p} + \theta \|g\|^2_{Sob,p} \tag{19}$$

subject to

$$-\mathbf{1} \leq \boldsymbol{f}'\left(\mathbb{x}_\alpha\right) \leq \mathbf{0},\ \mathbf{0} \leq \boldsymbol{g}'\left(\mathbb{x}_\beta\right) \leq \mathbf{1},\ \boldsymbol{f}''\left(\mathbb{x}_\alpha\right) \geq \mathbf{0},\ \boldsymbol{g}''\left(\mathbb{x}_\beta\right) \geq \mathbf{0},\ \boldsymbol{f}''\left(\mathbb{x}_\gamma\right) = \boldsymbol{g}''\left(\mathbb{x}_\gamma\right) \tag{20}$$

and

$$s^2 = \min_{\in \mathbb{R}^{\omega_Y},\ \in \mathbb{R}^{\omega_Z}} \left[\begin{pmatrix} \mathbb{Y} \\ \mathbb{Z} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\Delta} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Theta} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Psi} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Phi} \end{pmatrix} \begin{pmatrix} \mathbb{c} \\ \mathbb{d} \end{pmatrix} \right]^\top \boldsymbol{\Sigma}^{-1} \left[\begin{pmatrix} \mathbb{Y} \\ \mathbb{Z} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\Delta} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Theta} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Psi} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Phi} \end{pmatrix} \begin{pmatrix} \mathbb{c} \\ \mathbb{d} \end{pmatrix} \right] + \chi \mathbb{c}^\top \boldsymbol{\Psi} \mathbb{c} + \theta \mathbb{d}^\top \boldsymbol{\Phi} \mathbb{d} \tag{21}$$

subject to

$$-\mathbf{1} \leq \boldsymbol{\Psi}^{(1)} \mathbb{c} \leq \mathbf{0},\ \mathbf{0} \leq \boldsymbol{\Phi}^{(1)} \mathbb{d} \leq \mathbf{1},\ \boldsymbol{\Psi}^{(2)} \mathbb{c} \geq \mathbf{0},\ \boldsymbol{\Phi}^{(2)} \mathbb{d} \geq \mathbf{0},\ \boldsymbol{\Psi}^{(2)} \mathbb{c}_\gamma = \boldsymbol{\Phi}^{(2)} \mathbb{d}_\gamma, \tag{22}$$

where $\chi > 0, \theta > 0$, $\boldsymbol{\Sigma}$ *is* $(n+m) \times (n+m)$ *positive definite and symmetric covariance matrix,* $\boldsymbol{\Delta}$ *and* $\boldsymbol{\Theta}$ *are respectively the connectivity matrices for Call and Put options,* $\boldsymbol{\Psi}$ *is the* $\omega_Y \times \omega_Y$ *representor matrix at* $[x_\iota]^\top_{\iota \in \{\iota \,|\, n_\iota \geq 1\}}$, $\boldsymbol{\Phi}$ *is the* $\omega_Z \times \omega_Z$ *representor matrix at* $[x_\iota]^\top_{\iota \in \{\iota \,|\, m_\iota \geq 1\}}$, $\mathbb{Y} = [Y_1, \ldots, Y_n]^\top$, $\mathbb{Z} = [Z_1, \ldots, Z_m]^\top$, $\boldsymbol{f}(\mathbb{x}_\alpha) = [f(x_\iota)]^\top_{\iota \in \{\iota \,|\, n_\iota \geq 1\}}$, $\boldsymbol{g}(\mathbb{x}_\beta) = [g(x_\iota)]^\top_{\iota \in \{\iota \,|\, m_\iota \geq 1\}}$ *and* $\gamma := \alpha \cap \beta = [\iota \,|\, n_\iota \geq 1 \,\&\, m_\iota \geq 1]^\top$ *is the vector of indices in increasing order.*

Then $\hat{\sigma}^2 = s^2$. *Furthermore, there exists a solution to* (19) *with respect to* (20) *of the form*

$$\widehat{f} = \sum_{\{i \,|\, n_i \geq 1\}} \hat{c}_i \psi_{x_i} \quad \text{and} \quad \widehat{g} = \sum_{\{j \,|\, m_j \geq 1\}} \hat{d}_j \phi_{x_j}, \tag{23}$$

where $\widehat{\mathbb{c}} = [\hat{c}_i]^\top_{i \in \{i \,|\, n_i \geq 1\}}$ *and* $\widehat{\mathbb{d}} = [\hat{d}_j]^\top_{j \in \{j \,|\, m_j \geq 1\}}$ *solves* (21), ψ_{x_i} *is the representor at* x_i *for vector* $[x_\iota]^\top_{\iota \in \{\iota \,|\, n_\iota \geq 1\}}$ *and* ϕ_{x_j} *is the representor at* x_j *for vector* $[x_\iota]^\top_{\iota \in \{\iota \,|\, m_\iota \geq 1\}}$. *The estimators* $\widehat{f}$ *and* $\widehat{g}$ *are unique a.e.*

The structure of the $(n+m)\times(n+m)$ covariance matrix $\boldsymbol{\Sigma}$ of the random errors $(\varepsilon_1,\dots,\varepsilon_n,\nu_1,\dots,\nu_m)^\top$ will be investigated in Sect. 5. The minimization problem (21) under the constraints (22) can be implemented using, e.g., R statistical software with function `pcls()` in the library `mgcv`.

5 Covariance Structure

Let us denote the vector of the true SPD in the ω distinct observed strike prices $\varpi_1,\dots,\varpi_\omega$ as $h=[h(\varpi_1),\dots,h(\varpi_\omega)]^\top$. Assume that the expected values of the option prices given in (14) and (15) can be approximated by a linear combination of this discretized version of the SPD, i.e., we assume a linear model

$$Y_i=\alpha(x_{\alpha,i})^\top h+\varepsilon_i,\quad i=1,\dots,n\quad\text{and}\quad Z_j=\beta(x_{\beta,j})^\top h+\nu_j,\quad j=1,\dots,m$$

for the Call and Put option prices, respectively. We assume that the vectors of the coefficients $\alpha(x)$ and $\beta(x)$ depend only on the strike price x and can be interpreted as rows of design matrices $\mathscr{X}_\alpha$ and $\mathscr{X}_\beta$ so that the observed option prices can be written as

$$\begin{pmatrix}\mathbb{Y}\\ \mathbb{Z}\end{pmatrix}=\begin{pmatrix}\mathscr{X}_\alpha\\ \mathscr{X}_\beta\end{pmatrix}h+\begin{pmatrix}\boldsymbol{\varepsilon}\\ \boldsymbol{\nu}\end{pmatrix}. \tag{24}$$

In the following, the SPD may depend on the time of the observation and $h_k=[h_k(\varpi_1),\dots,h_k(\varpi_\omega)]^\top$ will denote the true value of the SPD at the time of the k-th trade, $k=1,\dots,n+m$. Such a time-ordering of the trades naturally orders the strike prices. Thus, the strike prices' *ranks* can be defined as follows: the strike price $x_{\alpha,i}$ of the i-th Call option price corresponding to the k-th trade has rank $r(\alpha,i)=k$. Similarly, the strike price $x_{\beta,j}$ of the j-th Call option price corresponding to the k-th trade has rank $r(\beta,j)=k$.

5.1 Constant SPD

Assuming that the random errors $\left(\boldsymbol{\varepsilon}^\top,\boldsymbol{\nu}^\top\right)=(\varepsilon_1,\dots,\varepsilon_n,\nu_1,\dots,\nu_m)$ in the linear model (24) are, the model (24) for the i-th observation, corresponding to the strike price $x_{\alpha,i}$, can be written as

$$Y_i=\alpha(x_{\alpha,i})^\top h_{r(\alpha,i)}+\varepsilon_i,\quad\text{where } h_{r(\alpha,i)}=h \text{ and } i=1,\dots,n, \tag{25}$$

for the i-th Call option price or

$$Z_j=\beta(x_{\beta,j})^\top h_{r(\beta,j)}+\nu_j,\quad\text{where } h_{r(\beta,j)}=h \text{ and } j=1,\dots,m, \tag{26}$$

for the j-th Put option price.

Here, the SPD $h = h_1 = \cdots = h_{n+m}$ is constant in the observation period. This simplified model has been investigated in [14] only for the Call option prices.

5.2 *Dependencies Due to the Time of the Trade*

Let us now assume that the observations are sorted according to the time of the trade $t_k \in (0, 1)$ and denote by $\delta_k = t_k - t_{k-1} > 0$ the time between the $(k-1)$-st and the k-th trade ($t_0 \equiv 0$). The models (25) and (26) can now be generalized by considering a different error structure

$$
\begin{aligned}
Y_i &= \alpha(x_{\alpha,i})^\top h_{r(\alpha,i)}, \quad i = 1, \ldots, n,\\
Z_j &= \beta(x_{\beta,j})^\top h_{r(\beta,j)}, \quad j = 1, \ldots, m,\\
h_k &= h_{k-1} + \delta_k^{1/2} \eta_k, \quad k = 1, \ldots, n+m,
\end{aligned}
$$

where $\eta_1, \ldots, \eta_{n+m}$ are independent and identically distributed random vectors with values in $\mathbb{R}^\omega$ having zero mean and variance matrix equal to $\varsigma^2 \mathbf{I}$ such that $\mathbf{I}$ is the $\omega \times \omega$ identity matrix. Expressing all observations in terms of an artificial parameter $h = h_0$, corresponding to the time 0 that can be interpreted as, e.g., "beginning of the day," it follows that the covariance of any two-observed call option prices depends only on their strike prices and on the time of the trade

$$
\begin{aligned}
\mathsf{Cov}(Y_u, Y_v) &= \mathsf{Cov}\left(\alpha(x_{\alpha,u})^\top h_{r(\alpha,u)}, \alpha(x_{\alpha,v})^\top h_{r(\alpha,v)}\right)\\
&= \varsigma^2 \alpha(x_{\alpha,u})^\top \alpha(x_{\alpha,v}) \sum_{s=1}^{\min\{r(\alpha,u), r(\alpha,v)\}} \delta_s.
\end{aligned} \tag{27}
$$

Similarly, we obtain the covariances between the observed Put option prices

$$
\begin{aligned}
\mathsf{Cov}(Z_u, Z_v) &= \mathsf{Cov}\left(\beta(x_{\beta,u})^\top h_{r(\beta,u)}, \beta(x_{\beta,v})^\top h_{r(\beta,v)}\right)\\
&= \varsigma^2 \beta(x_{\beta,u})^\top \beta(x_{\beta,v}) \sum_{s=1}^{\min\{r(\beta,u), r(\beta,v)\}} \delta_s.
\end{aligned} \tag{28}
$$

and the covariance between the observed Put and Call option prices

$$
\begin{aligned}
\mathsf{Cov}(Y_u, Z_v) &= \mathsf{Cov}\left(\alpha(x_{\alpha,u})^\top h_{r(\alpha,u)}, \beta(x_{\beta,v})^\top h_{r(\beta,v)}\right)\\
&= \varsigma^2 \alpha(x_{\alpha,u})^\top \beta(x_{\beta,v}) \sum_{s=1}^{\min\{r(\alpha,u), r(\beta,v)\}} \delta_s.
\end{aligned} \tag{29}
$$

Hence, the knowledge of the time of the trade allows us to approximate the covariance matrix of the observed option prices. Using this covariance structure, we can estimate arbitrary future value of the SPD. It is quite natural that more recent observations are more important for the construction of the estimator and that observations corresponding to the same strike price obtained at approximately same time will be highly correlated.

6 DAX Option Prices

In this section, the theory developed in the previous sections is applied on real data set consisting of intra-day Call and Put DAX option prices in year 1995. The data set, Eurex Deutsche Börse, was provided by the Financial and Economic Data Center (FEDC) at Humboldt-Universität zu Berlin in the framework of the SFB 649 Guest Researcher Program for Young Researchers.

From econometric theory, e.g., [14] or [6] with the discussions therein, follows that a practical choice is $p = 4$ in $\mathscr{H}^p$, which is used here as well. In Figs. 3 and 4, we present the analysis for the first two trading days in January 1995. On the first trading day, the time to expiry was $T - t = 0.05$ years, i.e., 18 days. Naturally, on the second trading day, the time to expiry was 17 days.

In both figures, the first two plots contain the fitted Put and Call option prices and the estimated SPD. Both smoothing parameters were chosen using the cross-validation as 2×10^{-5} leading to a reasonably smooth SPD estimate in the upper right plot in Figs. 3 and 4. Moreover, smaller values of the smoothing parameters would lead to a more variable and less smooth SPD estimates that would be difficult to interpret.

The second two plots in Figs. 3 and 4 show ordinary residual plots separately for the observed Put and Call option prices. The size of each plotting symbol denotes the number of residuals lying in the respective area. The shape of the plotting symbols corresponds to the time of the trade. The circles, squares and stars correspond, respectively, to morning, lunchtime, and afternoon prices. Clearly, we observe both heteroscedasticity and strong dependency due to the time of the trade.

In the last two plots in Figs. 3 and 4, we plot the same residuals transformed by Mahalanobis transformation, i.e., multiplied by the inverse square root of their assumed covariance matrix, see Sect. 5.2. This transformation removes most of the dependencies caused by the time of the trade. However, some outlying observations have now appeared. For example, for the Call options on the second day, plotted in Fig. 4, we can see a very large positive and a very large negative residual at the same strike price 2050.

The outlying observations can be explained if we have a closer look at the original data set. In Table 1, we show the Call option prices, times of the trades, and the transformed residuals for all trades with the strike price $K = 2050$. The two observations with large residuals, 358.7 and -342.2, occurred at approximately the same time, the time difference between them is approximately 0.13 h, i.e., approximately 5 minutes.

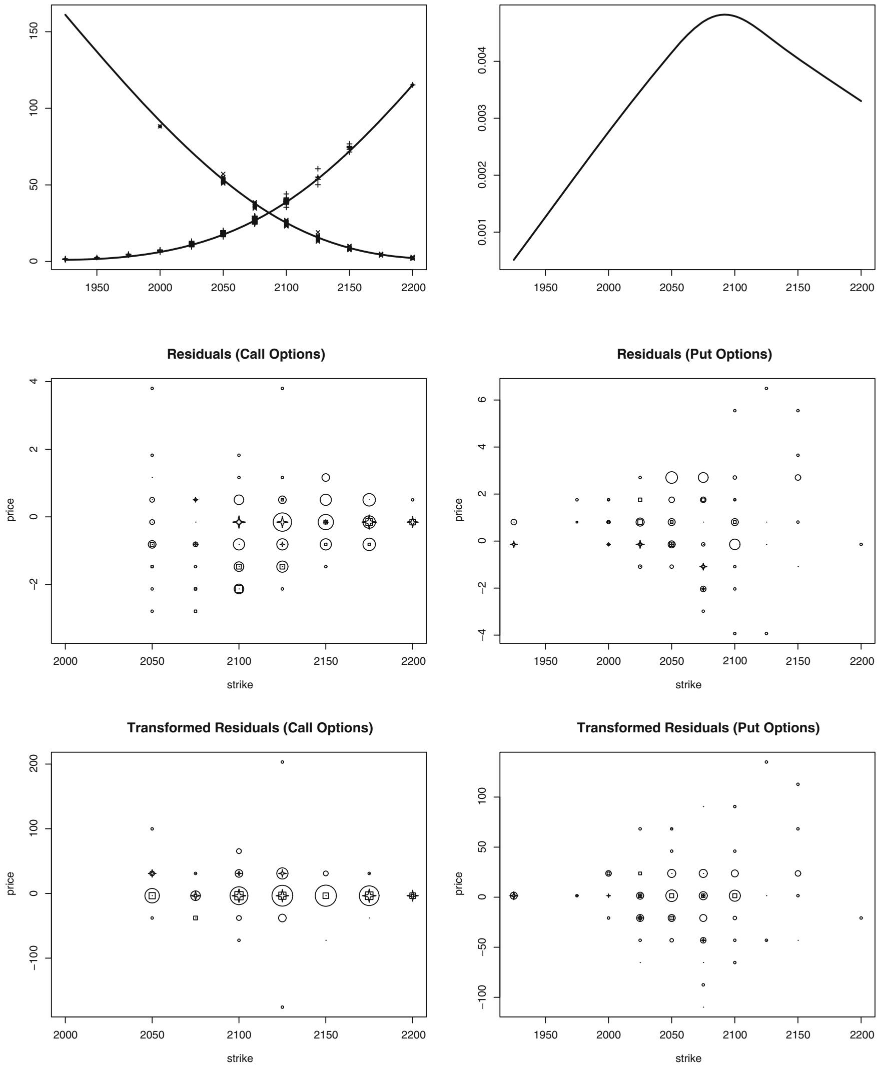


Fig. 3 Estimates and residual plots on the first trading day in 1995 (January 2). The *first plot* shows fitted Call and Put option prices, the estimated SPD is plotted in the *second plot*. The remaining four graphics contain, respectively, residual plots for Call and Put option prices on the left and right hand side. The residuals plotted in the last two plots were corrected by the inverse square root of the covariance matrix

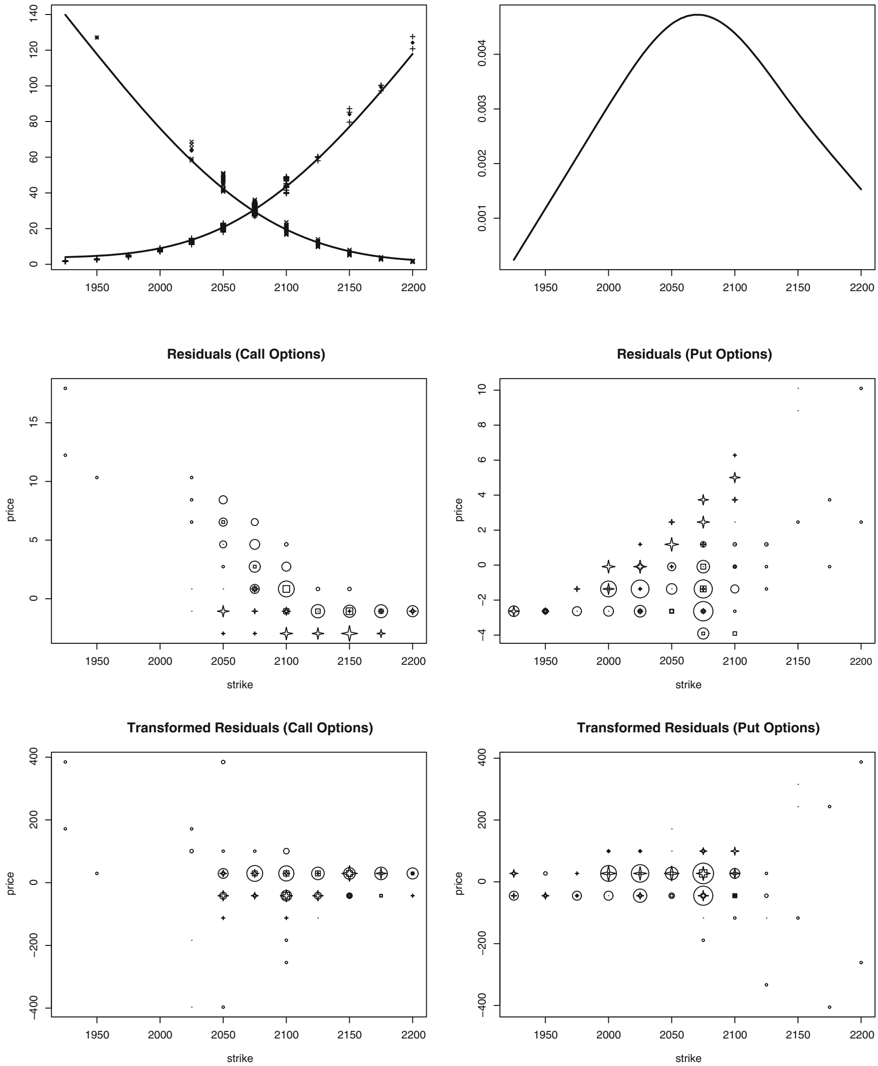


Fig. 4 Estimates and residual plots on second trading day in 1995 (January 3). The *first plot* shows fitted Call and Put option prices, the estimated SPD is plotted in the *second plot*. The remaining four graphics contain, respectively, residual plots for Call and Put option prices on the left- and right-hand side. The residuals plotted in the last two plots were corrected by the inverse square root of the covariance matrix

Table 1 Subset of observed prices of Call options on second trading day in 1995 for strike price $K = 2050$, time of the trade in hours and residuals transformed by the Mahalanobis transformation. The fitted value for the strike price $K = 2050$ is $\hat{f}^{(2)}(2050) = 42.37$. This value can be interpreted as an estimate corresponding to 16:00 o'clock

Call price ($K = 2050$)	Time (in hours)	Transformed residual
50.62296	9.690	337.4
51.12417	9.702	73.2
50.62296	9.785	33.8
50.02150	9.807	6.5
48.11687	9.826	−10.3
46.61322	9.864	−11.5
47.31492	10.121	−6.9
48.11687	10.171	26.5
49.01906	10.306	24.3
49.01906	10.361	26.3
50.32223	10.534	358.7
46.61322	10.666	−342.2
47.61565	10.672	32.8
45.00932	11.187	−62.2
48.11687	11.690	28.2
45.10957	12.100	−72.6
48.11687	12.647	53.9
48.11687	12.766	13.3
48.11687	13.170	28.3
47.51541	14.205	11.2
44.10713	14.791	−4.8
42.10226	15.137	−34.1
42.10226	15.138	−93.4
40.99958	15.232	−32.4
41.60104	15.250	−14.2
42.10226	15.283	−2.4
42.10226	15.288	−87.6
40.69885	15.638	−31.2
41.60104	15.658	−48.9
42.60348	15.711	−46.6
42.10226	15.715	6.7
41.60104	15.796	−39.2
42.10226	15.914	−49.5

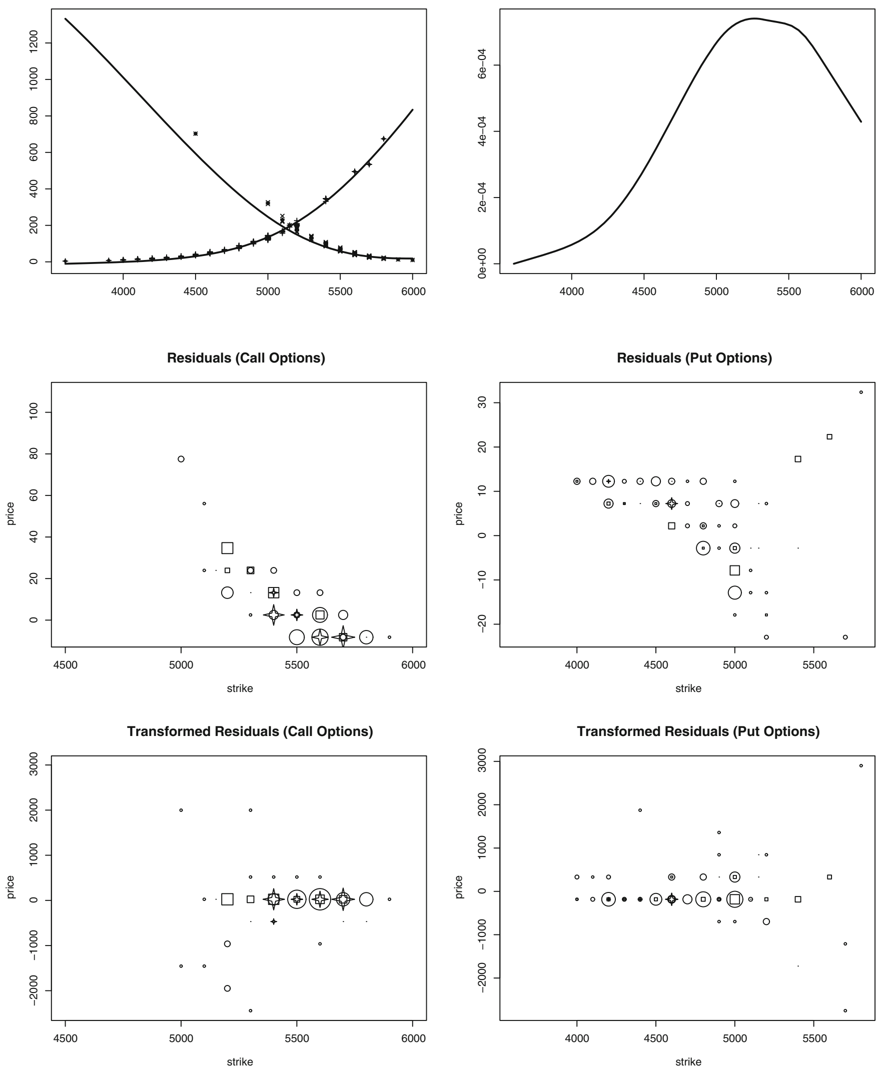


Fig. 5 Estimates and residual plots on the first trading day in 2002 (January 2). The *first plot* shows fitted Call and Put option prices, the estimated SPD is plotted in the *second plot*. The remaining four graphics contain respectively residual plots for Call and Put option prices on the left and right hand side. The residuals plotted in the last two plots were corrected by the inverse square root of the covariance matrix

Simultaneously, the price difference of these two observations is quite large. Hence, the large correlation of these two very different prices leads to the large (suspicious) residuals appearing in the residual plot.

An example of a more recent data set is plotted in Fig. 5. In year 2002, the range of the traded strike prices was much wider than in 1995. The estimated SPD is plotted in the upper right plot. The estimate could be described as a unimodal probability density function with the right tail cut off. It seems that, especially on the right hand side, the traded strike prices do not cover the entire support of the SPD.

The residual plots in Fig. 5 look very similar to the residual plots in Figs. 3 and 4. The residual analysis suggests that the simple model for the covariance structure presented in Sect. 5 is more appropriate for this estimation problem than the unrealistic iid assumptions. In practice, the traded strike prices do not cover the entire support of the SPD. Hence, our estimators recover only the central part of the SPD in Figs. 3 and 4 or the left hand part of the SPD in Fig. 5. Unfortunately, this implies that we cannot impose any conditions on the expected value of the SPD without additional distributional assumptions.

7 Conclusions

A nonparametric regression estimator for an unknown smooth regression curve with shape constraints is proposed. First, the penalized least squares are used in order to find a *compromise between the best fit and satisfactory smoothness*. Sobolev spaces allow us to transform the infinite-dimensional optimization problem into a finite-dimensional optimization. The estimator is derived in an alternative form compared to [13], more suitable from a *computational* point of view. We also consider *fixed design* that is more plausible for the data structure of observed option prices. We show that the estimator is consistent and establish its rate of convergence.

Second, *isotonic constraints* are placed on the estimator of the smooth regression curve in order to meet the requirements for the put–call option price duality. A covariance structure for the option prices, reflecting the time dependence, is suggested as well. Hence, an additional achievement of this paper is *simultaneous estimation* of the SPD from both put and call option prices and incorporation of the proposed *dependence structure* into the nonparametric estimator.

Acknowledgements The research of Michal Pešta was supported by the Czech Science Foundation project GAČR No. P201/13/12994P. Support through the Czech Science Foundation project GAČR No. 15-09663S is gratefully acknowledged by Zdeněk Hlávka.

Appendix: Proofs

Proof (*Proof of Theorem* 1) We divide the proof into two steps. The proof follows closely the proof of Theorem 2.2 given in [13]. However, we repeat it here since we need to introduce the notation needed for expressing the coefficients given in Theorem 2. This is, indeed, essential to derive the closed form of the estimator.

(i) *Construction of a representor* $\psi_a (\equiv \psi_a^0)$. For simplicity, let us set $\mathcal{Q}^1 \equiv [0, 1]$. For functions of one variable we have $\langle g, h\rangle_{Sob,m} = \sum_{k=0}^m \int_{\mathcal{Q}^1} g^{(k)}(x)h^{(k)}(x)\mathrm{d}x$. We are constructing a representor $\psi_a \in \mathcal{H}^m[0, 1]$ such that $\langle \psi_a, f\rangle_{Sob,m} = f(a)$ for all $f \in \mathcal{H}^m[0, 1]$. It suffices to demonstrate the result for all $f \in \mathcal{C}^{2m}$ because of the denseness of $\mathcal{C}^{2m}$. The representor is defined as

$$\psi_a(x) = \begin{cases} L_a(x) & 0 \le x \le a, \\ R_a(x) & a \le x \le 1, \end{cases} \tag{30}$$

where $L_a(x) \in \mathcal{C}^{2m}[0, a]$ and $R_a(x) \in \mathcal{C}^{2m}[a, 1]$. As $\psi_a \in \mathcal{H}^m[0, 1]$, it suffices that $L_a^{(k)}(a) = R_a^{(k)}(a)$, $0 \le k \le m-1$. We get

$$f(a) = \langle \psi_a, f\rangle_{Sob,m} = \int_0^a \sum_{k=0}^m L_a^{(k)}(x) f^{(k)}(x)\mathrm{d}x + \int_a^1 \sum_{k=0}^m R_a^{(k)}(x) f^{(k)}(x)\mathrm{d}x. \tag{31}$$

Integrating by parts and setting $i = k - j - 1$, we obtain

$$\begin{aligned} \sum_{k=0}^m \int_0^a L_a^{(k)}(x) f^{(k)}(x)\mathrm{d}x &= \sum_{i=0}^{m-1} f^{(i)}(a) \left\{ \sum_{k=i+1}^m (-1)^{k-i-1} L_a^{(2k-i-1)}(a) \right\} \\ - \sum_{i=0}^{m-1} f^{(i)}(0) &\left\{ \sum_{k=i+1}^m (-1)^{k-i-1} L_a^{(2k-i-1)}(0) \right\} + \int_0^a \left\{ \sum_{k=0}^m (-1)^k L_a^{(2k)}(x) \right\} f(x)\mathrm{d}x \end{aligned} \tag{32}$$

and, similarly,

$$\begin{aligned} \sum_{k=0}^m \int_a^1 R_a^{(k)}(x) f^{(k)}(x)\mathrm{d}x &= \sum_{i=0}^{m-1} f^{(i)}(1) \left\{ \sum_{k=i+1}^m (-1)^{k-i-1} R_a^{(2k-i-1)}(1) \right\} \\ - \sum_{i=0}^{m-1} f^{(i)}(a) &\left\{ \sum_{k=i+1}^m (-1)^{k-i-1} R_a^{(2k-i-1)}(a) \right\} + \int_a^1 \left\{ \sum_{k=0}^m (-1)^k R_a^{(2k)}(x) \right\} f(x)\mathrm{d}x. \end{aligned} \tag{33}$$

This holds for all $f(x) \in \mathcal{C}^m[0, 1]$. We require that both L_a and R_a are solutions of the constant coefficient differential equation

$$\sum_{k=0}^m (-1)^k \varphi_k^{(2k)}(x) = 0. \tag{34}$$

The boundary conditions are obtained by the equality of the functional values of $L_a^{(i)}(x)$ and $R_a^{(i)}(x)$ at a and the coefficient comparison of $f^{(i)}(0)$, $f^{(i)}(1)$ and $f^{(i)}(a)$, compare (31) to (32) and (33). Let $f^{(i)}(x) \bowtie c$ denotes that the term $f^{(i)}(x)$ has the coefficient c in a certain equation. We can write

$$r_a \in \mathscr{H}^m[0,1] \Rightarrow L_a^{(i)}(a) = R_a^{(i)}(a), \quad 0 \le i \le m-1, \tag{35}$$

$$f^{(i)}(0) \bowtie 0 \Rightarrow \sum_{k=i+1}^{m} (-1)^{k-i-1} L_a^{(2k-i-1)}(0) = 0, \quad 0 \le i \le m-1, \tag{36}$$

$$f^{(i)}(1) \bowtie 0 \Rightarrow \sum_{k=i+1}^{m} (-1)^{k-i-1} R_a^{(2k-i-1)}(1) = 0, \quad 0 \le i \le m-1, \tag{37}$$

$$f^{(i)}(a) \bowtie 0 \Rightarrow \sum_{k=i+1}^{m} (-1)^{k-i-1} \left\{L_a^{(2k-i-1)}(a) - R_a^{(2k-i-1)}(a)\right\} = 0, \quad 1 \le i \le m-1, \tag{38}$$

$$f(a) \bowtie 1 \Rightarrow \sum_{k=1}^{m} (-1)^{k-1} \left\{L_a^{(2k-1)}(a) - R_a^{(2k-1)}(a)\right\} = 1; \tag{39}$$

together $m + m + m + (m-1) + 1 = 4m$ boundary conditions. To obtain the general solution of this differential equation, we need to find the roots of its characteristic polynomial $P_m(\lambda) = \sum_{k=0}^{m}(-1)^k \lambda^{2k}$. Hence, it follows that

$$(1+\lambda^2)P_m(\lambda) = 1 + (-1)^m \lambda^{2m+2}, \quad \lambda \neq \pm i. \tag{40}$$

Solving (40), we get the characteristic roots $\lambda_k = e^{i\theta_k}$, where

$$\theta_k \in \begin{cases} \frac{(2k+1)\pi}{2m+2} & m \text{ even}, \; k \in \{0,1,\ldots,2m+1\} \setminus \left\{\frac{m}{2}, \frac{3m+2}{2}\right\}, \\ \frac{k\pi}{m+1} & m \text{ odd}, \; k \in \{0,1,\ldots,2m+1\} \setminus \left\{\frac{m+1}{2}, \frac{3m+3}{2}\right\}. \end{cases}$$

We have altogether $(2m+2) - 2 = 2m$ different complex roots but each has a pair that is conjugate with it. Thus, for $m$even, we have m complex conjugate roots with multiplicity one. We also have $2m$ base elements alike complex roots

$$\varphi_k(x) = \exp\left\{\left(\Re(\lambda_k)\right)x\right\} \cos\left[\left(\Im(\lambda_k)\right)x\right], \; k \in \{0,1,\ldots,m\} \setminus \left\{\frac{m}{2}\right\}; \tag{41}$$

$$\varphi_{m+1+k}(x) = \exp\left\{\left(\Re(\lambda_k)\right)x\right\} \sin\left[\left(\Im(\lambda_k)\right)x\right], \; k \in \{0,1,\ldots,m\} \setminus \left\{\frac{m}{2}\right\}. \tag{42}$$

If m is odd, we have $2m-2$ different complex roots (each has a pair that is conjugate with it) and two-real roots. The two real roots are ± 1. The $m-1$ complex conjugate roots have multiplicity one. We also have $2(m-1)+2 = 2m$ base elements alike all roots. These base elements are

$$\varphi_0(x) = \exp\{x\}; \tag{43}$$

$$\varphi_k(x) = \exp\left\{\left(\Re(\lambda_k)\right)x\right\}\cos\left[\left(\Im(\lambda_k)\right)x\right],\ k \in \{1, 2, \ldots, m\} \setminus \left\{\frac{m+1}{2}\right\}; \tag{44}$$

$$\varphi_{m+1}(x) = \exp\{-x\}; \tag{45}$$

$$\varphi_{m+1+k}(x) = \exp\left\{\left(\Re(\lambda_k)\right)x\right\}\sin\left[\left(\Im(\lambda_k)\right)x\right],\ k \in \{1, 2, \ldots, m\} \setminus \left\{\frac{m+1}{2}\right\}. \tag{46}$$

These vectors generate the subspace of $\mathscr{C}^m[0, 1]$ of solutions of the differential equation (34). The general solution is given by the linear combination

$$L_a(x) = \sum_{\substack{k=0 \\ k \neq \frac{m}{2}}}^{m} \gamma_k(a)\exp\left\{\Re(\lambda_k)x\right\}\cos\left[\Im(\lambda_k)x\right] + \sum_{\substack{k=0 \\ k \neq \frac{m}{2}}}^{m} \gamma_{m+1+k}(a)\exp\left\{\Re(\lambda_k)x\right\}\sin\left[\Im(\lambda_k)x\right], \quad \text{for } m \text{ even}; \tag{47}$$

$$L_a(x) = \gamma_0(a)\exp\{x\} + \sum_{\substack{k=1 \\ k \neq \frac{m+1}{2}}}^{m} \gamma_k(a)\exp\left\{\Re(\lambda_k)x\right\}\cos\left[\Im(\lambda_k)x\right] + \gamma_{m+1}(a)\exp\{-x\} + \sum_{\substack{k=1 \\ k \neq \frac{m+1}{2}}}^{m} \gamma_{m+1+k}(a)\exp\left\{\Re(\lambda_k)x\right\}\sin\left[\Im(\lambda_k)x\right], \quad \text{for } m \text{ odd}; \tag{48}$$

$$R_a(x) = \sum_{\substack{k=0 \\ k \neq \frac{m}{2}}}^{m} \gamma_{2m+2+k}(a)\exp\left\{\Re(\lambda_k)x\right\}\cos\left[\Im(\lambda_k)x\right] + \sum_{\substack{k=0 \\ k \neq \frac{m}{2}}}^{m} \gamma_{3m+3+k}(a)\exp\left\{\Re(\lambda_k)x\right\}\sin\left[\Im(\lambda_k)x\right], \quad \text{for } m \text{ even}; \tag{49}$$

$$R_a(x) = \gamma_{2m+2}(a)\exp\{x\} + \sum_{\substack{k=1 \\ k \neq \frac{m+1}{2}}}^{m} \gamma_{2m+2+k}(a)\exp\left\{\Re(\lambda_k)x\right\}\cos\left[\Im(\lambda_k)x\right] + \gamma_{3m+3}(a)\exp\{-x\} + \sum_{\substack{k=1 \\ k \neq \frac{m+1}{2}}}^{m} \gamma_{3m+3+k}(a)\exp\left\{\Re(\lambda_k)x\right\}\sin\left[\Im(\lambda_k)x\right], \quad \text{for } m \text{ odd}; \tag{50}$$

where the coefficients $\gamma_k(a)$ are arbitrary constants that satisfy the boundary conditions (35)–(39). It can be easily seen that we have obtained $4(m+1)-4=4m$ coefficients $\gamma_k(a)$, because the first index of $\gamma_k(a)$ is 0 and the last one is $4m+3$. Thus, we have $4m$ boundary conditions and $4m$ unknowns of γ_ks that lead us to the square $4m \times 4m$ system of the linear equations. Does ψ_a exist and is it unique? To show this, it suffices to prove that the only solution of the associated homogeneous system of linear equations is the zero vector. Suppose $L_a(x)$ and $R_a(x)$ are functions corresponding to the solution of the homogeneous system, because in linear system of equations (35)–(39) the right side has all zeros—coefficient of $f(a)$ in the last boundary condition is 0 instead of 1. Then, by the exactly the same integration by parts, it follows that $\langle \psi_a, f\rangle_{Sob,m} = 0$ for all $f \in \mathscr{C}^m[0,1]$. Hence, $\psi_a(x)$, $L_a(x)$ and $R_a(x)$ are zero almost everywhere and, by the linear independence of the base elements $\varphi_k(x)$, we obtain the uniqueness of the coefficients $\gamma_k(a)$.

(ii) *Producing a representor* ψ . Let us define the representor ψ by setting $\psi\ (\mathbb{x}) = \prod_{i=1}^{q} \psi_{a_i}^{w_i}(x_i)$ for all $\mathbb{x} \in \mathscr{Q}^q$, where $\psi_{a_i}^{w_i}(x_i)$ is the representor at a_i in $\mathscr{H}^m\left(Q^1\right)$. We know that $\mathscr{C}^m$ is dense in $\mathscr{H}^m$, so it is sufficient to show the result for $f \in \mathscr{C}^m(\mathscr{Q}^q)$. For simplicity let us suppose $\mathscr{Q}^q \equiv [0,1]^q$. After rewriting the inner product and using Fubini theorem we have

$$
\begin{aligned}
\langle \psi\ , f\rangle_{Sob,m} &= \left\langle \prod_{i=1}^{q} \psi_{a_i}^{w_i}, f \right\rangle_{Sob,m} \\
&= \sum_{i_1=0}^{m} \int_0^1 \frac{\partial^{i_1} \psi_{a_1}^{w_1}(x_1)}{\partial x_1^{i_1}} \left[\cdots \left[\sum_{i_q=0}^{m} \int_0^1 \frac{\partial^{i_q} \psi_{a_q}^{w_q}(x_q)}{\partial x_q^{i_q}} \cdot \frac{\partial^{i_1,\ldots,i_q} f(\mathbb{x})}{\partial x_1^{i_1} \ldots \partial x_q^{i_q}} \mathrm{d}x_q \right] \ldots \right] \mathrm{d}x_1.
\end{aligned}
$$

According to Definition 3 and notation in Definition 1, we can rewrite the center most bracket

$$
\begin{aligned}
\sum_{i_q=0}^{m} \int_0^1 \frac{\partial^{i_q} \psi_{a_q}^{w_q}(x_q)}{\partial x_q^{i_q}} \cdot \frac{\partial^{i_1,\ldots,i_q} f(\mathbb{x})}{\partial x_1^{i_1} \ldots \partial x_q^{i_q}} \mathrm{d}x_q &= \left\langle \psi_{a_q}^{w_q}, D^{(i_1,\ldots,i_{q-1})} f(x_1,\ldots,x_{i-1},\cdot) \right\rangle_{Sob,m} \\
&= D^{(i_1,\ldots,i_{q-1},w_q)} f(\mathbb{x}_{-q}, a_q).
\end{aligned}
$$

Proceeding sequentially in the same way, we obtain that the value of the above expression is $D\ \ f(\mathbb{a})$. □

Proof (*Proof of Theorem* 2) Existence and uniqueness of coefficients $\gamma_k(a)$ has already been proved in the proof of Theorem 1. Let us define

$$
\Lambda_{a,I}^{(l)} := \begin{cases} L_a^{(l)}(0), & \text{for } I = L; \\ R_a^{(l)}(1), & \text{for } I = R; \\ L_a^{(l)}(a) - R_a^{(a)}(a), & \text{for } I = D. \end{cases} \tag{51}
$$

From the boundary conditions (36)–(39), we easily see that

$$\sum_{k=i+1}^{m}(-1)^{k-i-1}\Lambda_{a,I}^{(2k-i-1)}=0,\ 0\leq i\leq m-1,\ I\in\{L,R,D\},\ [i,I]\neq[0,D]; \quad (52)$$

$$\sum_{k=1}^{m}(-1)^{k-1}\Lambda_{a,D}^{(2k-1)}=1. \quad (53)$$

For $m=1$ it follows from (52) to (53) that $\Lambda_{a,I}^{(1)}=0,\ I\in\{L,R\}$ and $\Lambda_{a,D}^{(1)}=1$. For $m=2$, we have from (52) to (53): $\Lambda_{a,I}^{(2)}=0,\ \forall I$; $\Lambda_{a,I}^{(1)}-\Lambda_{a,I}^{(3)}=0,\ I\in\{L,R\}$; and $\Lambda_{a,D}^{(1)}-\Lambda_{a,D}^{(3)}=1$. Let us now suppose that $m\geq 3$. We would like to prove the next important step:

$$\Lambda_{a,I}^{(m-j)}+(-1)^{j}\Lambda_{a,I}^{(m+j)}=0, \quad j=0,\dots,m-2,\ \forall I, \quad (54)$$

$$\Lambda_{a,I}^{(1)}+(-1)^{m-1}\Lambda_{a,I}^{(2m-1)}=0, \quad I\in\{L,R\}, \quad (55)$$

$$\Lambda_{a,D}^{(1)}+(-1)^{m-1}\Lambda_{a,D}^{(2m-1)}=1, \quad (56)$$

where $j:=m-i-1$. For $j=0$, we obtain $i=m-1$ and (52) and (53) implies

$$\Lambda_{a,I}^{(m)}=0, \quad \forall I, \quad (57)$$

which is correct according to (54). Consider $j=1$ and thus $i=m-2$. In the same way we get

$$\Lambda_{a,I}^{(m-1)}-\Lambda_{a,I}^{(m+1)}=0, \quad \forall I. \quad (58)$$

For $j=2$ and thus $i=m-3$, we have $\Lambda_{a,I}^{(m-2)}-\Lambda_{a,I}^{(m)}+\Lambda_{a,I}^{(m+2)}=0,\ \forall I$, and we can use (57). For $j=3$ and thus $i=m-4$ we have $\Lambda_{a,I}^{(m-3)}-\Lambda_{a,I}^{(m-1)}+\Lambda_{a,I}^{(m+1)}-\Lambda_{a,I}^{(m+3)}=0,\ \forall I$, where we can apply (58). We can continue in this way until $j=m-1$. The last step ensures the correctness of (55) in case that $I\in\{L,R\}$, eventually (56) if $I=D$ instead of (54).

To finish the proof, we only need to keep in mind (35). From (35), it follows that $\Lambda_{a,D}^{(j)}=0,\ j\in\{0,\dots,m-1\}$. According to (54) for $I=D$ and (56), we further see

$$\Lambda_{a,D}^{(j)}=0, \quad j\in\{m+1,\dots,2m-2\};$$
$$\Lambda_{a,D}^{(2m-1)}=(-1)^{m-1}.$$

Altogether we have obtained the following system of $4m$ linear equations:

$$\Lambda_{a,L}^{(m-j)}+(-1)^{j}\Lambda_{a,L}^{(m+j)}=0, \quad j=0,\dots,m-1,$$
$$\Lambda_{a,R}^{(m-j)}+(-1)^{j}\Lambda_{a,R}^{(m+j)}=0, \quad j=0,\dots,m-1,$$
$$\Lambda_{a,D}^{(j)}=0, \quad j=0,\dots,2m-2,$$
$$\Lambda_{a,D}^{(2m-1)}=(-1)^{m-1},$$

which, after rewriting them using (51), (47)–(50) and (41)–(46), completes the proof. □

Proof (*Proof of Theorem* 3) Let $M = span\left\{\psi_{\mathbb{x}_i} : i = 1, \ldots, n\right\}$ and its orthogonal complement $M^{\perp} = \left\{h \in \mathscr{H}^m : \left\langle\psi_{\mathbb{x}_i}, h\right\rangle_{Sob,m} = 0,\ i = 1, \ldots, n\right\}$. Representors exist by Theorem 1 and we can write the Sobolev space as a direct sum of its orthogonal subspaces, i.e., $\mathscr{H}^m = M \oplus M^{\perp}$ since $\mathscr{H}^m$ is a Hilbert space. Functions $h \in M^{\perp}$ take on the value zero at $\mathbb{x}_1, \ldots, \mathbb{x}_n$. Each $f \in \mathscr{H}^m$ can be written as $f = \sum_{j=1}^n c_j \psi_{\mathbb{x}_j} + h,\ h \in M^{\perp}$. Then,

$$
\begin{aligned}
&\left[\mathbb{Y} - \boldsymbol{f}(\mathbb{x})\right]^{\top} \boldsymbol{\Sigma}^{-1} \left[\mathbb{Y} - \boldsymbol{f}(\mathbb{x})\right] + \chi \left\|f\right\|^2_{Sob,m} \\
&= \left[\mathbb{Y}_{\bullet} - \left\langle \psi_{\mathbb{x}_\bullet}, \sum_{j=1}^n c_j \psi_{x_j} + h \right\rangle_{Sob,m}\right]^{\top} \boldsymbol{\Sigma}^{-1} \left[\mathbb{Y}_{\bullet} - \left\langle \psi_{\mathbb{x}_\bullet}, \sum_{j=1}^n c_j \psi_{x_j} + h \right\rangle_{Sob,m}\right] \\
&\quad + \chi \left\| \sum_{j=1}^n c_j \psi_{\mathbb{x}_j} + h \right\|^2_{Sob,m} \\
&= \left[\mathbb{Y}_{\bullet} - \sum_{j=1}^n \left\langle \psi_{\mathbb{x}_\bullet}, c_j \psi_{x_j} \right\rangle_{Sob,m}\right]^{\top} \boldsymbol{\Sigma}^{-1} \left[\mathbb{Y}_{\bullet} - \sum_{j=1}^n \left\langle \psi_{\mathbb{x}_\bullet}, c_j \psi_{x_j} \right\rangle_{Sob,m}\right] \\
&\quad + \chi \left\| \sum_{j=1}^n c_j \psi_{\mathbb{x}_j} \right\|^2_{Sob,m} + \chi \left\|h\right\|^2_{Sob,m} \\
&= \left[\mathbb{Y}_{\bullet} - \sum_{j=1}^n c_j \left\langle \psi_{\mathbb{x}_\bullet}, \psi_{x_j} \right\rangle_{Sob,m}\right]^{\top} \boldsymbol{\Sigma}^{-1} \left[\mathbb{Y}_{\bullet} - \sum_{j=1}^n c_j \left\langle \psi_{\mathbb{x}_\bullet}, \psi_{x_j} \right\rangle_{Sob,m}\right] \\
&\quad + \chi \left\langle \sum_{j=1}^n c_j \psi_{\mathbb{x}_j}, \sum_{j=1}^n c_j \psi_{\mathbb{x}_j} \right\rangle_{Sob,m} + \chi \left\|h\right\|^2_{Sob,m} \\
&= \left[\mathbb{Y}_{\bullet} - \sum_{j=1}^n \boldsymbol{\Psi}_{\bullet,j} c_j\right]^{\top} \boldsymbol{\Sigma}^{-1} \left[\mathbb{Y}_{\bullet} - \sum_{j=1}^n \boldsymbol{\Psi}_{\bullet,j} c_j\right] \\
&\quad + \chi \sum_{j=1}^n \sum_{k=1}^n c_j \left\langle \psi_{\mathbb{x}_j}, \psi_{\mathbb{x}_k} \right\rangle_{Sob,m} c_k + \chi \left\|h\right\|^2_{Sob,m} \\
&= \left[\mathbb{Y} - \boldsymbol{\Psi}\mathbb{c}\right]^{\top} \boldsymbol{\Sigma}^{-1} \left[\mathbb{Y} - \boldsymbol{\Psi}\mathbb{c}\right] + \chi \mathbb{c}^{\top} \boldsymbol{\Psi} \mathbb{c} + \chi \left\|h\right\|^2_{Sob,m},
\end{aligned}
$$

where $\left\langle\psi_{\mathbb{x}_\bullet}, g\right\rangle_{Sob,m} = \left[\left\langle\psi_{x_1}, g\right\rangle_{Sob,m}, \ldots, \left\langle\psi_{x_n}, g\right\rangle_{Sob,m}\right]^{\top}$ for an arbitrary $g \in \mathscr{H}^m$. Hence, there exists a function f^*, minimizing the infinite dimensional optimizing problem, that is a linear combination of the representors. We note also that $\|f^*\|^2_{Sob,m} = \mathbb{c}^\top \boldsymbol{\Psi} \mathbb{c}$.

Uniqueness is clear, since $\psi_{\ i}$ are the base elements of M, and adding a function that is orthogonal to the spaces spanned by the representors will increase the norm. □

Proof (*Proof of Corollary* 1) It follows directly from (30) and from Theorem 3. □

Proof (*Proof of Theorem* 4) The representor matrix is symmetric by Definition 4 since

$$\Psi_{i,j} = \langle \psi_{\ i}, \psi_{\ j} \rangle_{Sob,m} = \langle \psi_{\ j}, \psi_{\ i} \rangle_{Sob,m} = \Psi_{j,i},$$

i.e., $\boldsymbol{\Psi} = \boldsymbol{\Psi}^\top$.

We give the proof of positive definiteness of the representor matrix only for one-dimensional variable x. The extension into the multivariate case is straightforward. For an arbitrary $\mathbb{c} \in \mathbb{R}^n$, we obtain

$$\mathbb{c}^\top \boldsymbol{\Psi} \mathbb{c} = \sum_i c_i \sum_j \Psi_{ij} c_j = \sum_i \sum_j c_i \langle \psi_{x_i}, \psi_{x_j} \rangle_{Sob,m} c_j = \sum_i \sum_j \langle c_i \psi_{x_i}, c_j \psi_{x_j} \rangle_{Sob,m}$$
$$= \left\langle \sum_i c_i \psi_{x_i}, \sum_j c_j \psi_{x_j} \right\rangle_{Sob,m} = \left\| \sum_i c_i \psi_{x_i} \right\|^2_{Sob,m} \geq 0.$$

Hence, $\mathbb{c}^\top \boldsymbol{\Psi} \mathbb{c} = 0$ iff $\sum_i c_i \psi_{x_i} = 0$ a.e.

For $x > x_i$, we define

$$\boldsymbol{\gamma}(x_i) = \left[\gamma_0, \ldots, \gamma_{\kappa-1}, \gamma_{\kappa+1}, \ldots, \gamma_{m+\kappa}, \gamma_{m+2+\kappa}, \ldots, \gamma_{2m+1}\right]^\top (x_i).$$

Otherwise, we define

$$\boldsymbol{\gamma}(x_i) = \left[\gamma_{2m+2}, \ldots, \gamma_{2m+1+\kappa}, \gamma_{2m+3+\kappa}, \ldots, \gamma_{3m+2+\kappa}, \gamma_{3m+4+\kappa}, \ldots, \gamma_{4m+3}\right]^\top (x_i).$$

Similarly, we look at elements of the vector $\left[\{\boldsymbol{\Gamma}(x_i)\}^{-1}\right]_{\bullet,4m}$. According to (47)–(50), (41)–(46) and (7), we have

$$\psi_{x_i}(x) = \boldsymbol{\gamma}(x_i)^\top \boldsymbol{\varphi}(x) = (-1)^{m-1} \left[\{\boldsymbol{\Gamma}(x_i)\}^{-1}\right]^\top_{\bullet,4m} \boldsymbol{\varphi}(x)$$

where $\boldsymbol{\varphi}(x)$ is vector containing the base elements of the space of the solutions of the differential equation (34), i.e., $\varphi_k(x)$, cf. (41)–(46). From the linear independence of $\varphi_k(x)$ it follows that

$$\sum_i c_i \psi_{x_i} = (-1)^{m-1} \sum_i c_i \left[\{\boldsymbol{\Gamma}(x_i)\}^{-1} \right]^\top_{\bullet,4m} \boldsymbol{\varphi}$$
$$= (-1)^{m-1} \sum_i \sum_k c_i \left[\{\boldsymbol{\Gamma}(x_i)\}^{-1} \right]_{4m,k} \varphi_k = 0 \quad \text{a.e.}$$
$$\Updownarrow$$
$$\varphi_k = 0 \quad \text{a.e.} \quad k \in \{0, 1, \ldots, 2m+1\} \setminus \begin{cases} \left\{ \frac{m}{2}, \frac{3m+2}{2} \right\} & m \text{ even}, \\ \left\{ \frac{m+1}{2}, \frac{3m+3}{2} \right\} & m \text{ odd}; \end{cases}$$
$$\Downarrow$$
$$\psi_{x_i} = 0 \quad \text{a.e.} \quad i = 1, \ldots, n.$$

And $\psi_{x_i} = 0$ a.e. is a zero element of the space $\mathscr{H}^m$. □

Proof (*Proof of Theorem* 5) According to the Theorem 3, we want to minimize the function

$$\mathscr{L}(\mathbb{c}) := \frac{1}{n} [\mathbb{Y} - \boldsymbol{\Psi}\mathbb{c}]^\top \boldsymbol{\Sigma}^{-1} [\mathbb{Y} - \boldsymbol{\Psi}\mathbb{c}] + \chi \mathbb{c}^\top \boldsymbol{\Psi}\mathbb{c}.$$

Therefore, the first partial derivatives of $\mathscr{L}(\mathbb{c})$ have to be equal zero at the minimizer $\widehat{\mathbb{c}}$, i.e., $\frac{\partial}{\partial c_i}\mathscr{L}(\mathbb{c}) \stackrel{!}{=} 0,\ i = 1, \ldots, n$. Denoting $\boldsymbol{\Sigma}^{-1} =: \left[\phi_{ij}\right]_{i,j=1}^{n,n}$, we can write

$$n\mathscr{L}(\mathbb{c}) = \mathbb{Y}^\top \boldsymbol{\Sigma}^{-1}\mathbb{Y} - 2\mathbb{Y}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}\mathbb{c} + \mathbb{c}^\top \boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}\mathbb{c} + n\chi\mathbb{c}^\top\boldsymbol{\Psi}\mathbb{c}$$
$$= \sum_{r=1}^n \sum_{s=1}^n Y_r \phi_{rs} Y_s - 2\sum_{r=1}^n \sum_{s=1}^n \sum_{t=1}^n Y_r \phi_{rs} \Psi_{st} c_t + \sum_{r=1}^n \sum_{s=1}^n \sum_{t=1}^n \sum_{u=1}^n c_r \Psi_{rs} \phi_{st} \Phi_{tu} c_u$$
$$+ n\chi \sum_{r=1}^n \sum_{s=1}^n c_r \Psi_{rs} c_s$$

and, hence,

$$0 \stackrel{!}{=} -2\sum_{r=1}^n \sum_{s=1}^n Y_r \phi_{rs} \Psi_{si} + 2 \sum_{\substack{r=1 \\ r \neq i}}^n \sum_{s=1}^n \sum_{t=1}^n c_r \Psi_{rs} \phi_{st} \Phi_{ti}$$
$$+ 2\sum_{r=1}^n \sum_{s=1}^n c_i \Psi_{is} \phi_{st} \Phi_{ti} + 2n\chi \sum_{\substack{r=1 \\ r \neq i}}^n c_r \Psi_{ri} + 2n\chi c_i \Psi_{ii}$$
$$= -2\mathbb{Y}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Psi}_{\bullet,i} + 2\mathbb{c}^\top \boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1} \boldsymbol{\Psi}_{\bullet,i} + 2n\chi\mathbb{c}^\top \boldsymbol{\Psi}_{\bullet,i}, \quad i = 1, \ldots, n.$$

Then, we obtain our system of the normal equations

$$\mathbb{c}^\top \left(\boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\bullet,i} + n\chi\boldsymbol{\Psi}_{\bullet,i}\right) = \mathbb{Y}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\bullet,i}, \quad i = 1, \ldots, n.$$

□

Proof (*Proof of Theorem* 6) Let us have fixed $\chi > 0$. Hence we have obtained unique $\widehat{f}$ and also $\widehat{\mathbb{c}}$ according to Theorem 3. Theorem 3 and the Lagrange multipliers say that there exists a unique $L > 0$ such that $\widehat{\mathbb{c}}$ is also a unique solution of optimizing problem

$$\widehat{\mathbb{c}} = \arg\min_{\in\mathbb{R}^n} \frac{1}{n}\left[\mathbb{Y} - \boldsymbol{\Psi}\mathbb{c}\right]^\top \boldsymbol{\Sigma}^{-1}\left[\mathbb{Y} - \boldsymbol{\Psi}\mathbb{c}\right] \quad \text{s.t. } \mathbb{c}^\top\boldsymbol{\Psi}\mathbb{c} = L.$$

Let us define

$$\widetilde{\boldsymbol{f}}(\mathbb{x}) := \boldsymbol{\Xi}\boldsymbol{f}(\mathbb{x}), \quad \widetilde{\mathbb{Y}} := \boldsymbol{\Xi}\mathbb{Y},$$
$$\widehat{\widetilde{\mathbb{c}}} := \arg\min_{\widetilde{}\in\mathbb{R}^n} \frac{1}{n}\left[\widetilde{\mathbb{Y}} - \boldsymbol{\Psi}\widetilde{\mathbb{c}}\right]^\top \boldsymbol{\Sigma}^{-1}\left[\widetilde{\mathbb{Y}} - \boldsymbol{\Psi}\widetilde{\mathbb{c}}\right] \quad \text{s.t. } \widetilde{\mathbb{c}}^\top\boldsymbol{\Psi}\boldsymbol{\Xi}^{-1}\boldsymbol{\Psi}^{-1}\boldsymbol{\Xi}^{-1}\boldsymbol{\Psi}\widetilde{\mathbb{c}} \le L.$$

We can easily find out that $\widehat{\widetilde{\mathbb{c}}} = \boldsymbol{\Psi}^{-1}\boldsymbol{\Xi}\boldsymbol{\Psi}\widehat{\mathbb{c}}$ and, hence, $\widehat{\widetilde{\boldsymbol{f}}}(\mathbb{x}) = \boldsymbol{\Xi}\widehat{\widetilde{\mathbb{c}}}$. Finally, there must exists $\widetilde{L} > 0$ such that

$$\widehat{\widetilde{\mathbb{c}}} = \arg\min_{\widetilde{}\in\mathbb{R}^n} \frac{1}{n}\left[\widetilde{\mathbb{Y}} - \boldsymbol{\Psi}\widetilde{\mathbb{c}}\right]^\top \boldsymbol{\Sigma}^{-1}\left[\widetilde{\mathbb{Y}} - \boldsymbol{\Psi}\widetilde{\mathbb{c}}\right] \quad \text{s.t. } \widetilde{\mathbb{c}}^\top\boldsymbol{\Psi}\widetilde{\mathbb{c}} = \widetilde{L}$$

and hence this $\widehat{\widetilde{\mathbb{c}}}$ has to be a unique solution of the optimizing problem

$$\widehat{\widetilde{\mathbb{c}}} = \arg\min_{\widetilde{}\in\mathbb{R}^n} \frac{1}{n}\left[\widetilde{\mathbb{Y}} - \boldsymbol{\Psi}\widetilde{\mathbb{c}}\right]^\top \boldsymbol{\Sigma}^{-1}\left[\widetilde{\mathbb{Y}} - \boldsymbol{\Psi}\widetilde{\mathbb{c}}\right] \quad \text{s.t. } \widetilde{\mathbb{c}}^\top\boldsymbol{\Psi}\widetilde{\mathbb{c}} \le \widetilde{L}$$

since $\boldsymbol{\Psi}$ is a positive definite matrix ($\widetilde{\mathbb{c}}^\top\boldsymbol{\Psi}\widetilde{\mathbb{c}}$ is the volume of n-dimensional ellipsoid).

Now, we think of model

$$\widetilde{Y}_i = \widetilde{f}(\mathbb{x}_i) + \widetilde{\varepsilon}_i, \quad \widetilde{\varepsilon}_i \sim iid, \quad i = 1, \ldots, n$$

with least-squares estimator $\widehat{\widetilde{f}}$. As in the proof of Lemma 1 in [13], using [9], it can be shown that there exists $A > 0$ such that for $\delta > 0$, we have $\log N(\delta; \mathscr{F}) < A\delta^{-q/m}$, where $N(\delta; \mathscr{F})$ denotes the minimum number of balls of radius δ in sup-norm required to cover the set of functions $\mathscr{F}$. Consequently, applying [11, Lemma 3.5], we obtain that there exist positive constants C_0, K_0 such that for all $K > K_0$

$$\mathsf{P}\left[\sup_{\|g\|^2_{Sob,m} \le \widetilde{L}} \frac{\sqrt{n}\left|-\frac{2}{n}\sum_{i=1}^n \widetilde{\varepsilon}_i\left(\widetilde{f}(x_i) - g(x_i)\right)\right|}{\left(\frac{1}{n}\sum_{i=1}^n\left(\widetilde{f}(x_i) - g(x_i)\right)^2\right)^{\frac{1}{2}-\frac{q}{4m}}} \ge KA^{1/2}\right] \le \exp\left\{-C_0K^2\right\}.$$

Since $\widetilde{f} \in \widetilde{\mathscr{F}} = \left\{g \in \mathscr{H}^m(\mathscr{Q}^q) : \|g\|^2_{Sob,m} \leq \widetilde{L}\right\}$ and $\widehat{\widetilde{f}}$ minimizes the sum of squared residuals over $g \in \widetilde{\mathscr{F}}$,

$$\frac{1}{n}\sum_{i=1}^{n}\left[\widetilde{Y}_i - \widehat{\widetilde{f}}(x_i)\right]^2 \leq \frac{1}{n}\sum_{i=1}^{n}\left[\widetilde{Y}_i - g(x_i)\right]^2, \quad g \in \widetilde{\mathscr{F}}$$

$$\frac{1}{n}\sum_{i=1}^{n}\left[\left(\widetilde{f}(x_i) - \widehat{\widetilde{f}}(x_i)\right) + \widetilde{\varepsilon}_i\right]^2 \leq \frac{1}{n}\sum_{i=1}^{n}\left[\left(\widetilde{f}(x_i) - g(x_i)\right) + \widetilde{\varepsilon}_i\right]^2, \quad g \in \widetilde{\mathscr{F}}$$

$$\Downarrow \text{ realize that } \widetilde{f} \in \widetilde{\mathscr{F}}$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(\widetilde{f}(x_i) - \widehat{\widetilde{f}}(x_i)\right)^2 \leq -\frac{2}{n}\sum_{i=1}^{n}\widetilde{\varepsilon}_i\left(\widetilde{f}(x_i) - \widehat{\widetilde{f}}(x_i)\right). \tag{59}$$

Now combine (8) and (59) to obtain the result that $\forall K > K_0$

$$\mathsf{P}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\widetilde{f}(x_i) - \widehat{\widetilde{f}}(x_i)\right)^2 \geq \left(\frac{K^2 A}{n}\right)^{\frac{2m}{2m+q}}\right] \leq \exp\left\{-C_0 K^2\right\}.$$

Thus,

$$\frac{1}{n}\left[\widehat{\boldsymbol{f}}(\mathbb{x}) - \boldsymbol{f}(\mathbb{x})\right]^\top \boldsymbol{\Sigma}^{-1}\left[\widehat{\boldsymbol{f}}(\mathbb{x}) - \boldsymbol{f}(\mathbb{x})\right]$$
$$= \frac{1}{n}\sum_{i=1}^{n}\left(\widetilde{f}(x_i) - \widehat{\widetilde{f}}(x_i)\right)^2 = \mathscr{O}_{\mathsf{P}}\left(n^{-\frac{2m}{2m+q}}\right), \quad n \to \infty.$$

□

Lemma 1 *Suppose* $(f_n)_{n=1}^{\infty}$ *are nonnegative Lipschitz functions on interval* $[a, b]$ *with a constant* $T > 0$ *for all* $n \in \mathbb{N}$. *If* $f_n \xrightarrow[n\to\infty]{\mathsf{L}_1} 0$ *then*

$$\|f_n\|_{\infty,[a,b]} := \sup_{x\in[a,b]} |f_n(x)| \xrightarrow[n\to\infty]{} 0.$$

Proof (*Proof of Lemma* 1) Suppose that

$$\exists \rho > 0 \quad \forall n_0 \in \mathbb{N} \quad \exists n \geq n_0 \quad \exists x \in [a, b] \quad f_n(x) \geq \rho.$$

Then according to Lipschitz property of each $f_n \geq 0$, we have for fixed ρ, n_0, n and $x \in [a, b]$ that

$$\begin{aligned}
\|f_n\|_{\mathsf{L}_1[a,b]} &= \int_a^b f_n(t)\mathrm{d}t \\
&\geq \min\left\{\frac{f_n(x)}{2}(x-a)+\frac{f_n(x)}{2}(b-x), \frac{f_n(x)}{2}(x-a)+\frac{f_n(x)}{2}\frac{f_n(x)}{T},\right. \\
&\left.\frac{f_n(x)}{2}\frac{f_n(x)}{T}+\frac{f_n(x)}{2}(b-x), \frac{f_n(x)}{2}\frac{f_n(x)}{T}+\frac{f_n(x)}{2}\frac{f_n(x)}{T}\right\} \\
&\geq \min\left\{\frac{\rho}{2}(b-a), \frac{\rho}{2}(x-a)+\frac{\rho^2}{2T}, \frac{\rho^2}{2T}+\frac{\rho}{2}(b-x), \frac{\rho^2}{T}\right\} =: K > 0.
\end{aligned}$$

But K is a positive constant which does not depend on n and its existence would contradict the assumptions of this lemma, i.e.,

$$\forall \delta > 0 \quad \exists n_1 \in \mathbb{N} \quad \forall n \geq n_1 \quad \|f_n\|_{\mathsf{L}_1[a,b]} < \delta.$$

□

Proof (*Proof of Theorem* 7) We divide the proof into two steps.

(i) $s = 0$. The covariance matrix $\boldsymbol{\Sigma}$ is symmetric and positive definite with equibounded eigenvalues for all n. Hence it can be decomposed using Schur decomposition: $\boldsymbol{\Sigma} = \boldsymbol{\Gamma}\boldsymbol{\Upsilon}\boldsymbol{\Gamma}^\top$, where $\boldsymbol{\Gamma}$ is orthogonal, $\boldsymbol{\Upsilon}$ is diagonal (with eigenvalues on this diagonal) such that $0 < \Upsilon_{ii} \leq \vartheta$, $i = 1, \ldots, n$, $\forall n$. Hence $\boldsymbol{\Sigma}^{-1} = \boldsymbol{\Gamma} diag\left\{\Upsilon_1^{-1}, \ldots, \Upsilon_n^{-1}\right\}\boldsymbol{\Gamma}^\top$. Then

$$\begin{aligned}
&\frac{1}{n}\left[\widehat{\boldsymbol{f}}(\mathbb{x})-\boldsymbol{f}(\mathbb{x})\right]^\top \boldsymbol{\Sigma}^{-1}\left[\widehat{\boldsymbol{f}}(\mathbb{x})-\boldsymbol{f}(\mathbb{x})\right] \\
&\geq \frac{1}{n}\left[\widehat{\boldsymbol{f}}(\mathbb{x})-\boldsymbol{f}(\mathbb{x})\right]^\top \boldsymbol{\Gamma}\vartheta^{-1}\boldsymbol{I}\boldsymbol{\Gamma}^\top\left[\widehat{\boldsymbol{f}}(\mathbb{x})-\boldsymbol{f}(\mathbb{x})\right] = \frac{1}{n\vartheta}\sum_{i=1}^n\left[\hat{f}(x_i)-f(x_i)\right]^2.
\end{aligned} \tag{60}$$

Let us define $h_n := \left|\hat{f} - f\right|$. We know $\left\|\hat{f}\right\|^2_{Sob,m} \leq L$ for all n and $\left\|f\right\|^2_{Sob,m} \leq L$. For every function $t \in \mathscr{H}^m[a, b]$ with $\left\|t\right\|^2_{Sob,m} \leq L$, it holds that

$$\left\|t'\right\|_{\mathsf{L}_2[a,b]} \leq \|t\|_{Sob,1} \leq \|t\|_{Sob,m} \leq \sqrt{L}. \tag{61}$$

Then, t has equibounded derivative and hence there exists a Lipschitz constant $T > 0$ such that

$$|t(\xi) - t(\zeta)| < T\,|\xi - \zeta|\,, \quad \xi, \zeta \in [a, b].$$

We easily see

$$
\begin{aligned}
\frac{|h_n(\xi) - h_n(\zeta)|}{|\xi - \zeta|} &= \frac{\left| \left|\hat{f}(\xi) - f(\xi)\right| - \left|\hat{f}(\zeta) - f(\zeta)\right| \right|}{|\xi - \zeta|} \\
&\le \frac{\left| \left[\hat{f}(\xi) - f(\xi)\right] - \left[\hat{f}(\zeta) - f(\zeta)\right] \right|}{|\xi - \zeta|} \\
&\le \frac{\left|\hat{f}(\xi) - \hat{f}(\zeta)\right| + \left|f(\xi) - f(\zeta)\right|}{|\xi - \zeta|} < 2T, \quad \xi, \zeta \in [a, b].
\end{aligned}
$$

Since h_n is T-Lipschitz function for all n and

$$
\begin{aligned}
\|h_n\|_{\mathsf{L}_2[a,b]} = \left\|\hat{f} - f\right\|_{\mathsf{L}_2[a,b]} \le \left\|\hat{f} - f\right\|_{Sob,1} \le \left\|\hat{f} - f\right\|_{Sob,m} \\
\le \left\|\hat{f}\right\|_{Sob,m} + \left\|f\right\|_{Sob,m} \le 2\sqrt{L}, \quad \forall n,
\end{aligned}
$$

we obtain that h_n is equibounded for all n with a positive constant M such that $\|h_n\|_{\infty,[a,b]} \le M > 0,\ \forall n$. Hence, h_n^2 is also a Lipschitz function for all n, because for $\xi, \zeta \in [a, b]$

$$
\begin{aligned}
\frac{\left|h_n^2(\xi) - h_n^2(\zeta)\right|}{|\xi - \zeta|} = \frac{|h_n(\xi) - h_n(\zeta)|}{|\xi - \zeta|} \left[h_n(\xi) + h_n(\zeta)\right] \\
\le T \times 2 \|h_n\|_{\infty,[a,b]} = 2MT =: U > 0, \quad \forall n.
\end{aligned}
$$

Since h_n^2 is U-Lipschitz function for all n and design points $(x_i)_{i=1}^n$ are equidistantly distributed on $[a, b]$, we can write that

$$
\begin{aligned}
\int_a^b h_n^2(u)\mathrm{d}u &\le \sum_{i=1}^{n-1} \frac{x_{i+1} - x_i}{2} \left\{h_n^2(x_i) + \left[h_n^2(x_i) + U(x_{i+1} - x_i)\right]\right\} \\
&\le \frac{1}{2n}\left[2\sum_{i=1}^{n-1} h_n^2(x_i) + U(b - a)\right] \le \frac{1}{n}\sum_{i=1}^{n} h_n^2(x_i) + \frac{U(b - a)}{2n}.
\end{aligned} \tag{62}
$$

According to Theorem 6,

$$
\forall \rho > 0 \quad \mathsf{P}\left\{\frac{1}{n}\left[\widehat{\boldsymbol{f}}(\mathbb{x}) - \boldsymbol{f}(\mathbb{x})\right]^\top \boldsymbol{\Sigma}^{-1}\left[\widehat{\boldsymbol{f}}(\mathbb{x}) - \boldsymbol{f}(\mathbb{x})\right] > \rho\right\} \xrightarrow[n\to\infty]{} 0,
$$

so it means

$$
\begin{aligned}
&\forall \rho > 0 \quad \forall \delta > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \ge n_0 : \\
&\qquad \mathsf{P}\left\{\frac{1}{n}\left[\widehat{\boldsymbol{f}}(\mathbb{x}) - \boldsymbol{f}(\mathbb{x})\right]^\top \boldsymbol{\Sigma}^{-1}\left[\widehat{\boldsymbol{f}}(\mathbb{x}) - \boldsymbol{f}(\mathbb{x})\right] > \rho\right\} < \delta.
\end{aligned} \tag{63}
$$

Let us fix an arbitrary $\rho > 0$ and $\delta > 0$. Next, we fix $n_0 := \left\lceil \frac{U}{\rho^2} \right\rceil$ and for all $n \geq n_0$ we can write

$$\begin{aligned}
\delta &> \mathsf{P}\left\{ \frac{1}{n} \left[\widehat{\boldsymbol{f}}(\mathbb{x}) - \boldsymbol{f}(\mathbb{x})\right]^\top \boldsymbol{\Sigma}^{-1} \left[\widehat{\boldsymbol{f}}(\mathbb{x}) - \boldsymbol{f}(\mathbb{x})\right] > \frac{\rho^2(b-a)}{2\vartheta} \right\} && \text{by (63)} \\
&\geq \mathsf{P}\left\{ \frac{1}{n} \sum_{i=1}^{n} \left[\hat{f}(x_i) - f(x_i)\right]^2 > \frac{\rho^2(b-a)}{2} \right\} && \text{by (60)} \\
&\geq \mathsf{P}\left\{ \|h_n\|^2_{\mathsf{L}_2[a,b]} > \underbrace{\frac{\rho^2(b-a)}{2} + \frac{U(b-a)}{2n}}_{\tilde{\rho}} \right\} && \text{by (62)} \\
&\geq \mathsf{P}\left\{ \|h_n\|_{\mathsf{L}_1[a,b]} > \frac{\sqrt{\tilde{\rho}}}{\|1\|_{\mathsf{L}_2[a,b]}} \right\} \geq \mathsf{P}\left\{ \|h_n\|_{\mathsf{L}_1[a,b]} > \rho \right\} && \text{Cauchy–Schwarz.}
\end{aligned}$$

Thus, $\|h_n\|_{\mathsf{L}_1[a,b]} \xrightarrow[n\to\infty]{\mathsf{P}} 0$. According to Lemma 1 and the fact that the almost sure convergence implies convergence in probability, we have

$$\sup_{x\in[a,b]} \left|\widehat{f}(x) - f(x)\right| \xrightarrow[n\to\infty]{\mathsf{P}} 0.$$

(ii) $s \geq 1$. If $m = 2$, we are done. Let $g_n := \widehat{f} - f$. According to the assumptions of our model, $g_n \in \mathscr{H}^m[a,b]$. By [13, Theorem 2.3], all functions in the estimating set have derivatives up to order $m-1$ uniformly bounded in sup-norm. Then, all the g_n'' are also bounded in sup-norm ($m \geq 3$) and this implies the uniform boundedness of g_n'':

$$\exists M > 0 \quad \forall n \in \mathbb{N} \quad \left\|g_n''\right\|_{\infty,[a,b]} < M.$$

Let us have fixed $M > 0$. For any fixed $\rho > 0$, define $\tilde{\rho} := M\rho$ and there exists $n_0 \in \mathbb{N}$, such that $\forall n \geq n_0 : [c_n, d_n] \subset [a,b]$ and

$$g_n'(c_n) = g_n'(d_n) = \tilde{\rho} \quad \& \quad g_n'(\xi) > \tilde{\rho},\ \xi \in (c_n, d_n)$$

because g_n' is continuous on $[c_n, d_n]$ (drawing a picture is helpful). If such $[c_n, d_n]$ does not exist, the proof is finished.

Otherwise there exists $n_1 \geq n_0$ such that $\forall n \geq n_1$ holds

$$|\tilde{\rho}(d_n - c_n)| \leq \left|\int_{c_n}^{d_n} g_n'(\xi)\mathrm{d}\xi\right| = |g_n(d_n) - g_n(c_n)| \leq 2\rho^2$$

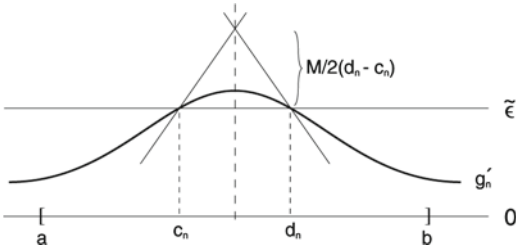


Fig. 6 Uniform convergence of g_n'

because $g_n \xrightarrow{n\to\infty} 0$ uniformly in sup-norm on the interval $[a, b]$. Hence, $|d_n - c_n| \leq \frac{2\rho}{M}$. The uniform boundedness of g_n'' implies Lipschitz property (see Fig. 6):

$$\left|g_n'(x)\right| \leq \left|\tilde{\rho} + M\frac{d_n - c_n}{2}\right| \leq M\rho + M\frac{\rho}{M} \leq \rho(M+1).$$

We can continue in this way finitely times (formally we can proceed by something like a finite induction). In fact, if $(m-1)$-th derivatives are uniformly bounded ($g_n \in \mathscr{H}^m[a, b]$), then this ensures that $\widehat{f}^{(s)}$ for $s \leq m-2$ converges in sup-norm. Finally, we have to realize that convergence almost sure implies convergence in probability and each convergent sequence in probability has a subsequence that converges almost sure. □

Proof (*Proof of Theorem* 8) The proof is very similar to the proof of the Infinite to Finite Theorem 3 and the same arguments can be used. Each $f, g \in \mathscr{H}^m$ can be written in the form:

$$f = \sum_{\{i\,|\,n_i \geq 1\}} c_i \psi_{x_i} + h_f, \quad h_f \in \left\{span\left\{\psi_{\ i} : n_i \geq 1\right\}\right\}^{\perp},$$
$$g = \sum_{\{j\,|\,m_j \geq 1\}} d_j \phi_{x_j} + h_g, \quad h_g \in \left\{span\left\{\phi_{\ j} : m_j \geq 1\right\}\right\}^{\perp}.$$

For $1 \leq \iota \leq n$, we easily note that

$$\left[\begin{pmatrix}\mathbb{Y}\\ \mathbb{Z}\end{pmatrix} - \begin{pmatrix}\boldsymbol{\Delta} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\Theta}\end{pmatrix}\begin{pmatrix}\boldsymbol{f}\left(\mathbb{x}_\alpha\right)\\ \boldsymbol{g}\left(\mathbb{x}_\beta\right)\end{pmatrix}\right]_\iota = Y_\iota - \left\{\sum_{\{i\,|\,n_i \geq 1\}} \Delta_{\iota i} f(x_i) + \sum_{\{i\,|\,m_i \geq 1\}} \Theta_{\iota i} g(x_i)\right\}$$
$$= Y_\iota - \sum_{\{i\,|\,n_i \geq 1\}} \Delta_{\iota i} \left\langle \psi_{x_i}, \sum_{\{j\,|\,n_j \geq 1\}} c_j \psi_{x_j} + h_f \right\rangle_{Sob,m}$$
$$- \sum_{\{i\,|\,m_i \geq 1\}} \Theta_{\iota i} \left\langle \phi_{x_i}, \sum_{\{j\,|\,m_j \geq 1\}} d_j \phi_{x_j} + h_g \right\rangle_{Sob,m}$$

$$= Y_\iota - \sum_{\{i \mid n_i \ge 1\}} \Delta_{\iota i} \sum_{\{j \mid n_j \ge 1\}} \Psi_{ij} c_j - \sum_{\{i \mid m_i \ge 1\}} \Theta_{\iota i} \sum_{\{j \mid m_j \ge 1\}} \Phi_{ij} d_j$$

$$= \left[\begin{pmatrix} \mathbb{Y} \\ \mathbb{Z} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\Delta} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Theta} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Psi} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Phi} \end{pmatrix} \begin{pmatrix} \mathbb{c} \\ \mathbb{d} \end{pmatrix} \right]_\iota .$$

We can proceed in the same way also for $n < \iota \le n + m$.

Finally, it remains to rewrite the constraints using (2) from Theorem 1:

$$f'(x_\iota) = \left\langle \psi_{x_\iota}, \sum_{\{i \mid n_i \ge 1\}} c_i \psi'_{x_i} + h_f \right\rangle_{Sob,m} = \left[\boldsymbol{\Psi}^{(1)} \mathbb{c} \right]_\iota \quad \forall \iota : n_\iota \ge 1.$$

Similarly, we obtain $g'(x_\iota) = \left[\boldsymbol{\Phi}^{(1)} \mathbb{d} \right]_\iota \ \forall \iota : m_\iota \ge 1$; $f''(x_\iota) = \left[\boldsymbol{\Psi}^{(2)} \mathbb{c} \right]_\iota \ \forall \iota : n_\iota \ge 1$; and $g''(x_\iota) = \left[\boldsymbol{\Phi}^{(2)} \mathbb{d} \right]_\iota \ \forall \iota : m_\iota \ge 1$. □

References

1. Aït-Sahalia, Y., Lo, A.W.: Nonparametric risk management and implied risk aversion. J. Econ. **94**(1–2), 9–51 (2000)
2. Aït-Sahalia, Y., Wang, Y., Yared, F.: Do option markets correctly price the probabilities of movement of the underlying asset? J. Econ. **102**(1), 67–110 (2001)
3. Barlow, R., Bartholomew, D., Bremner, J., Brunk, H.: Statistical Inference Under Order Restrictions: The Theory and Application of Isotonic Regression. Wiley, New York (1972)
4. Breeden, D.T., Litzenberger, R.H.: Prices of state-contingent claims implicit in option prices. J. Bus. **51**, 621–651 (1978)
5. Eubank, R.L.: Nonparametric Regression and Spline Smoothing. Marcel Dekker Inc, New York (1999)
6. Härdle, W., Hlávka, Z.: Dynamics of state price densities. J. Econ. **150**(1), 1–15 (2009)
7. Hlávka, Z., Svojík, M.: Application of extended Kalman filter to SPD estimation. In: Härdle, W.K., Hautsch, N., Overbeck, L. (eds.) Applied Quantitative Finance, pp. 233–247. Springer, Berlin (2008)
8. Jackwerth, J.C.: Option-implied risk-neutral distributions and implied binomial trees: a literature review. J. Deriv. **7**, 66–82 (1999)
9. Kolmogorov, A.N., Tihomirov, V.M.: ε-entropy and ε-capacity of sets in function spaces. Uspehi Matematičeskih Nauk **14**(2), 3–86 (1959). (English transl.: Am. Math. Soc. Transl.: Ser. 2. 17, 277–364 (1961))
10. Rao, C.: Linear Statistical Inference and Its Applications, 2nd edn. Wiley, New York (1973)
11. Van de Geer, S.: Estimating a regression function. Ann. Stat. **18**(2), 907–924 (1990)
12. Villalobos, M., Wahba, G.: Inequality-constrained multivariate smoothing splines with application to the estimation of posterior probabilities. J. Am. Stat. Assoc. **82**(397), 239–248 (1987)
13. Yatchew, A., Bos, L.: Nonparametric least squares estimation and testing of economic models. J. Quant. Econ. **13**, 81–131 (1997)
14. Yatchew, A., Härdle, W.: Nonparametric state price density estimation using constrained least squares and the bootstrap. J. Econ. **133**(2), 579–599 (2006)

On Existence of Explicit Asymptotically Normal Estimators in Nonlinear Regression Problems

Alexander Sakhanenko

Abstract Explicit asymptotically normal estimators for two new classes of nonlinear regression problems are constructed. The survey of such estimators and of methods for their construction is presented. Several new properties of previously established estimators are found.

Keywords Nonlinear regression · Explicit estimators · Asymptotically normal estimators · Improvement of estimators

1 Introduction

Let $Y_1, Y_2, \ldots, Y_n, \ldots$ be independent observations which may be represented in the following form:

$$Y_i = g_i(\theta) + \varepsilon_i \quad \text{with} \quad \mathbf{E}\varepsilon_i = 0 \quad \text{and} \quad 0 < \text{var } \varepsilon_i < \infty, \tag{1}$$

for all $i = 1, 2, \ldots,$ where $\{g_i(\theta)\}$ are known functions, whereas random variables $\{\varepsilon_i\}$ are independent and unobservable. Note that quite often one has $g_i(\theta) = g(x_i, \theta)$, where $\{x_i\}$ are known numbers or vectors.

Our aim is to estimate the unknown parameter $\theta \in \mathbf{R}$ or $\theta = (\theta_1, \ldots, \theta_k) \in \mathbf{R}^k$ using only the first n observations. If for all $i = 1, 2, \ldots$

$$\mathbf{E}\varepsilon_i = 0 \quad \text{and} \quad 0 < \text{var } \varepsilon_i = \sigma^2/w_{oi} < \infty, \tag{2}$$

where $\{w_{oi}\}$ are known positive numbers, then the standard advice is to use the famous Weighted Least Square Estimator $\hat{\theta}_n$ which may be defined as a solution of the following equation:

This work is supported by the RFBR-grant #15-01-07460a.

A. Sakhanenko (✉)
Sobolev Institute of Mathematics, Koptuga prospekt 4, Novosibirsk 630090, Russia
e-mail: aisakh@mail.ru

J. Antoch et al. (eds.), *Analytical Methods in Statistics*, Springer Proceedings in Mathematics & Statistics 193, DOI 10.1007/978-3-319-51313-3_8

$$\sum_{i=1}^{n} w_{oi}(Y_i - g_i(\hat{\theta}_n))^2 = \min_{\theta} \sum_{i=1}^{n} w_{oi}(Y_i - g_i(\theta))^2. \tag{3}$$

In case of the linear regression the famous solution $\hat{\theta}_n$ of Eq. (3) may be found in an explicit way as the solution of a system of linear equations. In particular, if the parameter θ is one-dimensional and

$$Y_i = \theta x_i + \varepsilon_i, \qquad i = 1, 2, \ldots,$$

then

$$\hat{\theta}_n = \sum_{i=1}^{n} w_{oi} x_i Y_i \Big/ \sum_{i=1}^{n} w_{oi} x_i^2 = \theta + \sum_{i=1}^{n} w_{oi} x_i \varepsilon_i \Big/ \sum_{i=1}^{n} w_{oi} x_i^2. \tag{4}$$

Thus, if random variables $\{w_{oi} x_i \varepsilon_i\}$ satisfy Lindeberg condition then the estimator $\hat{\theta}_n$ is asymptotically normal, i.e.,

$$(\hat{\theta}_n - \theta_n)/D_n \Rightarrow \mathcal{N}(0, 1) \qquad \text{as} \quad n \to \infty, \tag{5}$$

where the asymptotic variance D_n^2 has the following form:

$$D_n^2 = \frac{\sigma^2}{\sum_{i=1}^{n} w_{io} x_i^2} = \frac{1}{\sum_{i=1}^{n} (g_i'(\theta))^2 \big/ \operatorname{var} \ \varepsilon_i} \to 0. \tag{6}$$

Simple Explicit Estimators

The natural conjecture may arise that the explicit asymptotically normal estimators exist only in the linear regression problems. But in 2000 in [1], it was shown that if

$$Y_i = \frac{a_i}{1 + b_i \theta} + \varepsilon_i \qquad \text{for each} \quad i = 1, 2, \ldots,$$

where $a_i > 0$ and $b_i > 0$ are known numbers, then the following estimator

$$\theta_n^* = \sum_{i=1}^{n} c_i (a_i - Y_i) \Big/ \sum_{i=1}^{n} c_i b_i Y_i \tag{7}$$

is asymptotically normal (see details in Sect. 2) for a large class of constants $\{c_i\}$. That is,

$$(\theta_n^* - \theta)/d_n \Rightarrow \mathcal{N}(0, 1) \qquad \text{as} \quad n \to \infty$$

for some asymptotic variance $d_n^2 \to 0$. We would like to stress that the estimators in (7) are not Weighted Least Square Estimators no matter which weights $\{w_{oi}\}$ are chosen.

Furthermore, note that even in the classical case when $\{w_{oi} = 1\}$, the search for the Least Square Estimators could present a substantial computational burden since potentially there could be a growing random number of local extremes of the function in the right-hand side of the Eq. (3). The monograph [2] contains a good survey of the issues that arise and ways to overcome them. There one can also find several examples of seemingly simple regression problems for which it is very complicated to find the Least Square Estimators.

However, while considering these examples in [2] it occurred to the author of this paper to try to find sufficiently exact explicit asymptotically normal estimators using different methods and get analogous results as in [1]. Three such examples of estimators were found in 2013–2015 in papers [3–5] in cases when

$$Y_i = x_i\theta + z_i g(\theta) + \varepsilon_i, \quad i = 1, 2, \ldots, \tag{8}$$

where numbers $\{x_i\}$ and $\{z_i\}$ are known (see details in Sect. 3); also in cases when

$$Y_i = \sqrt{1 + x_i\theta} + \delta_i, \quad i = 1, 2, \ldots \tag{9}$$

(see Sect. 4 for details); and, at last, in cases (see Sect. 5) when we have independent observations

$$Y_i = \ln(1 + x_i\theta) + \delta_i, \qquad i = 1, 2, \ldots. \tag{10}$$

Let us stress that in (9) and (10) and later on all $\{x_i\}$ are given numbers and we denote by $\{\delta_i\}$ a sequence of i.i.d. unobservable random variables (with some additional properties on their common distribution, of course).

In the present paper, we construct explicit asymptotically normal estimators for two new classes of regression problems when

$$Y_i = (1 + x_i\theta)^r + \varepsilon_i = (1 + x_i\theta)^r(1 + \delta_i), \quad i = 1, 2, \ldots, \tag{11}$$

where $r \neq 0$ is a known number (see Sect. 6), and when

$$Y_i = e^{x_i\theta} + \varepsilon_i = e^{x_i\theta}(1 + \delta_i), \quad i = 1, 2, \ldots \tag{12}$$

(see details in Sect. 7).

In Sects. 9 and 10, several multidimensional explicit estimators are also considered.

Larger Classes of Explicit Estimators

Note, that all explicit estimators θ_n^* constructed in Sects. 2–7, 9, and 10 are elementary functions of finite number of linear statistics $\sum_{i=1}^n h_{ki}(Y_i)$ with specially chosen functions $\{h_{ki}(\cdot)\}$. But if we constructed a "sufficiently good" initial estimator θ_n^*, then we may introduce a larger class of explicit estimators θ_n^{**} that may be represented as functions of a finite number of statistics $\sum_{i=1}^n H_{ki}(Y_i, \theta_n^*)$ with specially chosen functions $\{H_{ki}(\cdot, \cdot)\}$.

Suppose now that for all $i = 1, 2, \ldots$

$$\mathbf{E}\varepsilon_i = 0 \quad \text{and} \quad 0 < \text{var } \varepsilon_i = \sigma^2 / w_i(\theta) < \infty, \tag{13}$$

where $0 < \sigma < \infty$ may be unknown. Introduce

$$D_{n,o}^2 := \frac{\sigma^2}{\sum_{i=1}^n w_i(\theta)(g_i'(\theta))^2} = \frac{1}{\sum_{i=1}^n (g_i'(\theta))^2 \big/ \text{var } \varepsilon_i}. \tag{14}$$

One of our aims in one-dimensional case is to find such asymptotically normal explicit estimator θ_n^{**} that

$$(\theta_n^{**} - \theta) \big/ D_{n,o} \Rightarrow \mathcal{N}(0, 1) \quad \text{as} \quad n \to \infty. \tag{15}$$

Our interest of finding this kind of estimators is motivated by the observation that $D_{n,o}^2$ is the minimal asymptotic variance for various well-known classes of statistical estimators. We will discuss this point later in Sect. 8. For example, it follows from (5) and (6) that the simplest optimal estimator $\hat{\theta}_n$ from (4) satisfy (15) with the minimal asymptotic variance $D_n^2 = D_{n,o}^2$ introduced in (14).

Everywhere in the paper limits are taken with respect to $n \to \infty$.

2 Fractional-Linear Regression

Suppose we have independent observations

$$Y_i = \frac{a_i}{1 + b_i\theta} + \varepsilon_i, \qquad i = 1, 2, \ldots, \tag{16}$$

where $\{a_i > 0\}$ and $\{b_i > 0\}$ are given numbers. We assume that independent variables $\{\varepsilon_i\}$ may have different distributions with

$$\mathbf{E}\varepsilon_i = 0 \quad \text{and} \quad 0 < \sigma_i^2 := \text{var } \varepsilon_i < \infty, \quad i = 1, 2, \ldots. \tag{17}$$

We want to construct an explicit estimator for the unknown parameter $\theta > 0$.

Construction of Estimators

First of all, we rewrite (16) in the following way:

$$(1 + b_i\theta)Y_i = a_i + (1 + b_i\theta)\varepsilon_i, \qquad \text{or} \qquad a_i - Y_i = \theta b_i Y_i + \tilde{\varepsilon}_i, \tag{18}$$

where

$$\tilde{\varepsilon}_i = -(1 + b_i\theta)\varepsilon_i, \qquad \mathbf{E}\tilde{\varepsilon}_i = 0, \qquad \text{var } \tilde{\varepsilon}_i = (1 + b_i\theta)^2 \text{ var } \varepsilon_i. \tag{19}$$

Now, for any constants $\{c_i\}$ we obtain from (18) that

$$\sum_{i=1}^{n} c_i(a_i - Y_i) = \theta \sum_{i=1}^{n} c_i b_i Y_i + \sum_{i=1}^{n} c_i \tilde{\varepsilon}_i.$$

So, we have the following representation:

$$\theta = \sum_{i=1}^{n} c_i(a_i - Y_i) \Big/ \sum_{i=1}^{n} c_i b_i Y_i - \sum_{i=1}^{n} c_i \tilde{\varepsilon}_i \Big/ \sum_{i=1}^{n} c_i b_i Y_i. \tag{20}$$

It follows from (20) that it is natural to define an estimator θ_n^* in the following way:

$$\theta_n^* := \sum_{i=1}^{n} c_i(a_i - Y_i) \Big/ \sum_{i=1}^{n} c_i b_i Y_i. \tag{21}$$

Asymptotic Normality of Estimators

It follows from (20) and (21) that

$$\theta_n^* - \theta = \frac{\sum_{i=1}^{n} c_i \tilde{\varepsilon}_i}{\sum_{i=1}^{n} c_i b_i Y_i} = \frac{\sum_{i=1}^{n} c_i \tilde{\varepsilon}_i}{\sum_{i=1}^{n} c_i a_i \beta_i + \sum_{i=1}^{n} c_i \beta_i \tilde{\varepsilon}_i}, \tag{22}$$

where $\beta_i := b_i/(1 + b_i\theta) < 1/\theta$. Assume that

$$d_n^2(\{c_i\}) := \sum\nolimits_{i=1}^{n} c_i^2 \operatorname{var} \tilde{\varepsilon}_i \Big/ \Big(\sum\nolimits_{i=1}^{n} c_i a_i \beta_i\Big)^2 \to 0. \tag{23}$$

It is not difficult to derive the following assertion from the representation (22).

Theorem 1 *Suppose that independent $\{c_i\tilde{\varepsilon}_i\}$ satisfy Lindeberg condition and that (19) and (23) hold. Then the statistic θ_n^* from (21) is an asymptotically normal estimator for θ:*

$$(\theta_n^* - \theta)/d_n(\{c_i\}) \Rightarrow \mathcal{N}(0, 1). \tag{24}$$

Corollary 1 *Suppose that $\{\varepsilon_i/\sigma_i\}$ are i.i.d. and for some constants $\{c_i\}$*

$$\inf_i \min\{a_i, b_i, c_i, \sigma_i\} > 0, \qquad \sup_i \max\{a_i, b_i, c_i, \sigma_i\} < \infty.$$

Then (24) holds with $d_n(\{c_i\}) = O(1/\sqrt{n})$.

Improved Estimators

Together with the preliminary estimators (22), let us introduce a larger class of estimators

$$\theta_n^{**} = \sum_{i=1}^{n} \gamma_i(\theta_n^*)(a_i - Y_i) \Big/ \sum_{i=1}^{n} \gamma_i(\theta_n^*) b_i Y_i. \tag{25}$$

Theorem 2 *Under assumptions of Corollary 1 suppose that all functions* $\{\gamma_i(t)\}$ *have derivatives* $\{\gamma_i'(t)\}$ *and*

$$\inf_i \gamma_i(\theta) > 0, \quad \sup_i \gamma_i(\theta) < \infty, \quad \sup_i \sup_{\theta/2 \le t \le 2\theta} |\gamma_i'(t)| < \infty.$$

Then

$$(\theta_n^{**} - \theta)/d_n(\{\gamma_i(\theta)\}) \Rightarrow \mathcal{N}(0, 1).$$

Optimization of Estimators

Theorem 3 *For each* $n \ge 1$

$$\inf_{\{c_i\}} d_n^2(\{c_i\}) = \inf_{\{\gamma_i(\theta)\}} d_n^2(\{\gamma_i(\theta)\}) = \frac{1}{\sum_{i=1}^n a_i^2 \beta_i^2 / \operatorname{var} \tilde{\varepsilon}_i} = d_n^2\big(\{\gamma_{o,i}(\theta)\}\big),$$

where

$$\gamma_{o,i}(\theta) = \gamma_{o,i}(\theta, \operatorname{var} \varepsilon_i, C) = C \frac{a_i \beta_i}{\operatorname{var} \tilde{\varepsilon}_i} = C \frac{a_i b_i}{(1 + b_i \theta)^3 \operatorname{var} \varepsilon_i} \quad \forall\, C \ne 0.$$

Corollary 2 *If* $\operatorname{var} \varepsilon_i = \sigma^2 / w_i(\theta)$ *for all* i*, then*

$$d_{n,opt}^2 = \frac{\sigma^2}{\sum_{i=1}^n a_i^2 b_i^2 w_i(\theta)/(1 + b_i\theta)^4} \quad \textit{with} \quad \gamma_{opt,i}(\theta) = \frac{a_i b_i w_i(\theta)}{(1 + b_i \theta)^3}. \tag{26}$$

Corollary 2 follows from Theorem 3 with $C = \sigma^2$.

Partial Cases

Example 1 If conditions (2) hold then

$$\gamma_{opt,i}(\theta) = a_i b_i w_{oi}/(1 + b_i\theta)^3.$$

In this case, we may recommend to take $c_i = \gamma_{opt,i}(\theta_0)$ for some $\theta_0 \ge 0$.

Example 2 If (17) holds with

$$\mathbf{E}\varepsilon_i = 0 \quad \text{and} \quad 0 < \operatorname{var} \varepsilon_i = \frac{\sigma^2}{w_{oi}(1 + b_i\theta)^3}, \quad i = 1, 2, \ldots \tag{27}$$

then we have $\gamma_{opt,i}(\theta) = a_i b_i w_{oi}$. So, we do not need to improve the estimator θ_n^* in this case if only we take $c_i = a_i b_i w_{oi}$.

Example 3 The simplest choice for c_i is to take $c_i = a_i$. But this choice is optimal only in the case when (27) holds with $w_{oi} = 1/b_i$.

Note that the simple Theorem 1 is a new result. All other assertions in this section may be easily derived from the corresponding results in [1].

3 Partially Linear Regression

We have independent observations

$$Y_i = x_i\theta + z_i g(\theta) + \varepsilon_i, \quad i = 1, 2, \ldots, \tag{28}$$

where numbers $\{x_i\}$ and $\{z_i\}$ are known, and independent variables $\{\varepsilon_i\}$ may have different distributions satisfying (17). We do not assume that the function $g(\cdot)$ is known. Our aim is to estimate the unknown parameter θ.

Construction of Explicit Estimators

First of all, we need to choose constants $\{c_{ni}\}$ such that

$$\sum_{i=1}^{n} c_{ni} z_i = 0 \quad \text{and} \quad \sum_{i=1}^{n} c_{ni} x_i \neq 0 \quad \forall\, n \geq n_0 \tag{29}$$

for some $n_0 < \infty$. It follows from (28) and (29) that

$$\sum_{i=1}^{n} c_{ni} Y_i = \theta \sum_{i=1}^{n} c_{ni} x_i + 0 \cdot g(\theta) + \sum_{i=1}^{n} c_{ni} \varepsilon_i .$$

So that

$$\theta = \sum_{i=1}^{n} c_{ni} Y_i \Big/ \sum_{i=1}^{n} c_{ni} x_i - \sum_{i=1}^{n} c_{ni} \varepsilon_i \Big/ \sum_{i=1}^{n} c_{ni} x_i . \tag{30}$$

It is natural from (30) to introduce an estimator θ_n^* in the following way:

$$\theta_n^* = \theta_n^*(c_{n\bullet}) = \sum_{i=1}^{n} c_{ni} Y_i \Big/ \sum_{i=1}^{n} c_{ni} x_i . \tag{31}$$

Asymptotic Normality of Estimators

It follows from (30) and (31) that

$$\theta_n^*(c_{n\bullet}) - \theta = \sum_{i=1}^{n} c_{ni} \varepsilon_i \Big/ \sum_{i=1}^{n} c_{ni} x_i .$$

Thus, we are ready to prove asymptotic normality.

Theorem 4 (See [3]) *Suppose that independent* $\{c_{ni}\varepsilon_i\}$ *satisfy Lindeberg condition and that* (17) *and* (29) *hold. Then*

$$\frac{\theta_n^*(c_{n\bullet})-\theta}{d_n(c_{n\bullet})} \Rightarrow \mathcal{N}(0,1) \quad \textit{where} \quad d_n^2(c_{n\bullet}) := \mathrm{var}\ \theta_n^*(c_{n\bullet}) = \frac{\sum_{i=1}^n c_{ni}^2\sigma_i^2}{\left(\sum_{i=1}^n c_{ni}x_i\right)^2}.$$

Moreover, if independent random variables $\{\varepsilon_i\}$ *have normal distributions with zero means then* $(\theta_n^*(c_{n\bullet})-\theta)/d_n(c_{n\bullet})$ *has the standard normal distribution.*

Optimization of Estimators

If conditions (2) hold and $\sum_{i=1}^n w_{oi}z_i^2 > 0$ we may introduce numbers

$$c_{ni,opt} = w_{oi}(x_i - k_n z_i) \quad \text{with} \quad k_n = \frac{\sum_{i=1}^n w_{oi}z_i x_i}{\sum_{i=1}^n w_{oi}z_i^2}.$$

And we set $k_n = 0$ when $\sum_{i=1}^n w_{oi}z_i^2 = 0$. Note that if

$$K_n := \sum_{i=1}^n w_{oi}(x_i - k_n z_i)^2 > 0 \tag{32}$$

then numbers $\{c_{ni,opt}\}$ satisfy both conditions in (29).

Theorem 5 (See [3]) *Let conditions* (2) *and* (32) *hold for some* n. *Then for all constants* $\{c_{ni}\}$ *satisfying (29) we have the following property:*

$$\mathrm{var}\ \theta_n^*(c_{n\bullet}) \geq d_{n,opt}^2 := \mathrm{var}\,\theta_n^*(c_{n\bullet,opt}) = \sigma^2/K_n > 0.$$

Remark 1 If we have no information about the behavior of the unknown $\{\mathrm{var}\ \varepsilon_i\}$ we may advise to use numbers $\{c_{ni} = c_{ni,opt}\}$ with $\{w_{oi} = 1\}$ as in the case of i.i.d. $\{\varepsilon_i\}$.

Possible Generalizations

Remark 2 Note that in Theorem 5 we have found optimal asymptotic variances $d_{n,opt}^2$ only for the estimators of the form (31). If we know the function $g(\cdot)$, then we may improve θ_n^*, using ideas from Sect. 8. In this case, we may obtain asymptotically normal estimators with better asymptotic variances than in Theorem 5 (at least in the case when condition (13) holds).

Remark 3 Of course, generalizations to the case

$$Y_i = \sum_{j=1}^m x_{ji}\theta_j + \sum_{k=1}^m z_{ki}g_k(\theta) + \varepsilon_i \quad i = 1, 2, \ldots,$$

with unknown m-dimensional parameter $(\theta_1, \ldots, \theta_m)$, are also possible.

4 Power Regression of Order 1/2

Suppose we have independent observations

$$Y_i = \sqrt{1 + x_i\theta} + \delta_i, \quad i = 1, 2, \ldots, \tag{33}$$

where $\{x_i > 0\}$ are known numbers and $\{\delta_i\}$ are i.i.d. with

$$\mathbf{E}\delta_1 = 0, \quad \mathbf{E}\delta_1^4 < \infty, \quad \sigma^2 := \text{var } \delta_1 > 0. \tag{34}$$

Our aim is to construct an explicit estimator for the unknown parameter $\theta > 0$.

Construction of Estimators

Rewrite (33) as

$$Y_i^2 = 1 + \sigma^2 + x_i\theta + \tilde{\varepsilon}_i, \quad i = 1, 2, \ldots, \tag{35}$$

where

$$\tilde{\varepsilon}_i := 2\delta_i\sqrt{1 + x_i\theta} + \delta_i^2 - \sigma^2 = Y_i^2 - \mathbf{E}Y_i^2.$$

It is easy to see that

$$\mathbf{E}\tilde{\varepsilon}_i = 0, \quad \text{var } \tilde{\varepsilon}_i = 4\sigma^2(1 + x_i\theta) + 4\sqrt{1 + x_i\theta}\mathbf{E}\delta_1^3 + \mathbf{E}\delta_1^4 - \sigma^4. \tag{36}$$

Now, introduce constants $\{c_{ni}\}$ such that

$$\sum_{i=1}^n c_{ni} = 0 \quad \text{and} \quad \sum_{i=1}^n c_{ni}x_i \neq 0 \quad \forall\, n \geq n_0. \tag{37}$$

It is possible to do if $n_0 := \min\{n : x_n \neq x_1\} < \infty$. Multiplying (35) by $\{c_{ni}\}$ we obtain

$$\sum_{i=1}^n c_{ni}Y_i^2 = \sum_{i=1}^n (1 + \sigma^2)c_{ni} + \sum_{i=1}^n c_{ni}\theta x_i + \sum_{i=1}^n c_{ni}\tilde{\varepsilon}_i = \theta\sum_{i=1}^n c_{ni}x_i + \sum_{i=1}^n c_{ni}\tilde{\varepsilon}_i.$$

Thus, we have

$$\theta = \sum_{i=1}^n c_{ni}Y_i^2 \Big/ \sum_{i=1}^n c_{ni}x_i, - \sum_{i=1}^n c_{ni}\tilde{\varepsilon}_i \Big/ \sum_{i=1}^n c_{ni}x_i. \tag{38}$$

So, it is natural to introduce an estimator

$$\theta_n^* := \sum_{i=1}^n c_{ni}Y_i^2 \Big/ \sum_{i=1}^n c_{ni}x_i. \tag{39}$$

Asymptotic Normality of Estimators

It follows from (38) and (39) that

$$\theta_n^* - \theta = \sum_{i=1}^{n} c_{ni}\tilde{\varepsilon}_i \Big/ \sum_{i=1}^{n} c_{ni}x_i.$$

Thus, we are ready to investigate asymptotic normality.

Theorem 6 (See [4]) *Let independent random variables* $\{c_{ni}\tilde{\varepsilon}_i\}$ *satisfy Lindeberg condition and* (34), (37) *hold. Then*

$$\frac{\theta_n^* - \theta}{d_n(c_{n\bullet})} \Rightarrow \mathcal{N}(0,1), \quad where \quad d_n(c_{n\bullet}) = \frac{\sqrt{\sum_{i=1}^{n} c_{ni} \operatorname{var} \tilde{\varepsilon}_i}}{\sum_{i=1}^{n} c_{ni}x_i}. \tag{40}$$

The Simplest Partial Case

Here we set

$$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \overline{Y_n^2} = \frac{1}{n}\sum_{i=1}^{n} Y_i^2.$$

Theorem 7 (See [4]) *Suppose that* $\{\delta_i\}$ *are i.i.d. with*

$$\mathbf{E}\delta_1 = \mathbf{E}\delta_1^3 = 0, \quad \mathbf{E}\delta_1^4 < \infty, \quad \operatorname{var}\ \delta_1 > 0.$$

Also assume

$$\max_{i \le n} x_i^3 \Big/ \sum_{i=1}^{n}(x_i - \bar{x}_n)^2 \to 0.$$

Then (40) *is true for* $\{c_{ni} = x_i - \bar{x}_n\}$. *That is the following estimator*

$$\theta_n^* = \sum_{i=1}^{n}(x_i - \bar{x}_n)(Y_i^2 - \overline{Y_n^2}) \Big/ \sum_{i=1}^{n}(x_i - \bar{x}_n)^2. \tag{41}$$

is asymptotically normal.

Optimization and Generalizations

Remark 4 The transformed equation (35) may be rewritten in the form

$$\tilde{Y}_i := Y_i^2 - 1 = \sigma^2 + x_i\theta + \tilde{\varepsilon}_i, \quad i = 1, 2, \ldots.$$

So, it is the well-known equation of the linear regression with two unknown parameters. But this equation is heteroscedastic with the complicated formula (36) for var $\tilde{\varepsilon}_i$. Nevertheless, in [4] a formula for values $\{c_{ni} = c_{ni,opt}\big(x_i, \{\text{var}\ \tilde{\varepsilon}_i\}\big)\}$ was found which minimizes the asymptotic variance $d_n^2(c_{n\bullet})$ defined in (40). But there is no use of this formula because it essentially depends on the unknown parameters (var $\tilde{\varepsilon}_1, \ldots,$ var $\tilde{\varepsilon}_n$).

In such a situation, we advise to use the simplest estimator from (41) as preliminary and after that introduce an improved estimator described in Sect. 8. For example we may advise to use

$$\theta_n^{**} = \theta_n^* - 2\sum_{i=1}^n \left(x_i - \frac{x_i Y_i}{\sqrt{1 + x_i\theta_n^*}}\right) \Big/ \sum_{i=1}^n \frac{x_i^2}{1 + x_i\theta_n^*}.$$

A paper with properties of the estimator θ_n^{**} is now in preparation.

Remark 5 Of course, all these results may be extended to cases of a multidimensional parameter θ when

$$Y_i = \sqrt{x_{0i} + x_{1i}\theta_1 + \cdots + x_{mi}\theta_m} + \varepsilon_i, \quad i = 1, 2, \ldots,$$

where $(\theta_1, \ldots, \theta_m)$ is an m-dimensional unknown parameter.

5 Logarithmic Regression

We have observations

$$Y_i = \ln(1 + x_i\theta) + \delta_i, \qquad i = 1, 2, \ldots, \tag{42}$$

where $\{x_i > 0\}$ are known numbers. We assume that unobservable random variables $\delta_1, \delta_2, \ldots$ are i.i.d. with

$$0 < \text{var}\ e^{\delta_1} < \infty \quad \text{and} \quad n_0 := \min\{n : x_n \neq x_1\} < \infty.$$

Our aim is to construct a simple explicit estimator for the main unknown parameter $\theta > 0$ using the first $n \geq n_0$ observations from (42).

Construction of Explicit Estimators

Our idea is as follows. First of all, we rewrite (42):

$$e^{Y_i} = (1 + x_i\theta)e^{\delta_i}, \qquad i = 1, 2, \ldots. \tag{43}$$

Introduce notations:

$$Z_i := e^{Y_i}, \qquad \beta := \mathbf{E}e^{\delta_1}, \qquad \tilde{\delta}_i := e^{\delta_i}/\beta - 1.$$

Here we have introduced a new parameter β. It is easy to see that

$$\{\tilde{\delta}_i\} \text{ are i.i.d. with } \mathbf{E}\tilde{\delta}_1 = 0 \quad \text{and} \quad 0 < \tilde{\sigma}^2 := \text{var } \tilde{\delta}_1 < \infty. \tag{44}$$

Let us stress that Eqs. (42) and (43) are equivalent to the following one:

$$Z_i = \beta(1 + x_i\theta)(1 + \tilde{\delta}_i) = \beta + x_i\beta\theta + \tilde{\varepsilon}_i, \qquad i = 1, 2, \dots \tag{45}$$

with $\tilde{\varepsilon}_i = \beta(1 + x_i\theta)\tilde{\delta}_i$.

Next we choose constants $\{a_{ni}\}$ and $\{b_{ni}\}$ such that

$$\sum\nolimits_{i=1}^n a_{ni} = 0 \quad \text{and} \quad A_n := \sum\nolimits_{i=1}^n a_{ni}x_i \neq 0, \tag{46}$$

$$\sum\nolimits_{i=1}^n b_{ni}x_i = 0 \quad \text{and} \quad B_n := \sum\nolimits_{i=1}^n b_{ni} \neq 0. \tag{47}$$

Multiplying (45) by $\{a_{ni}\}$ and $\{b_{ni}\}$ we obtain

$$\sum_{i=1}^n a_{ni}Z_i = \sum_{i=1}^n a_{ni}((1 + x_i\theta)\beta + \sigma_i\eta_i) = \theta\beta A_n + \sum_{i=1}^n a_{ni}\tilde{\varepsilon}_i, \tag{48}$$

$$\sum_{i=1}^n b_{ni}Z_i = \sum_{i=1}^n b_{ni}((1 + x_i\theta)\beta + \sigma_i\eta_i) = \beta B_n + \sum_{i=1}^n b_{ni}\tilde{\varepsilon}_i. \tag{49}$$

Thus, we have from (48) and (49) that

$$\theta = \frac{\theta\beta}{\beta} = \frac{\sum_{i=1}^n a_{ni}Z_i/A_n - \sum_{i=1}^n a_{ni}\tilde{\varepsilon}_i/A_n}{\sum_{i=1}^n b_{ni}Z_i/B_n - \sum_{i=1}^n b_{ni}\tilde{\varepsilon}_i/B_n}. \tag{50}$$

So, it is natural to introduce the estimator θ^* in the following way:

$$\theta_n^* = \theta_n^*(a_{n\bullet}, b_{n\bullet}) := \frac{\sum_{i=1}^n a_{ni}Z_i/A_n}{\sum_{i=1}^n b_{ni}Z_i/B_n} = \frac{B_n\sum_{i=1}^n a_{ni}Z_i}{A_n\sum_{i=1}^n b_{ni}Z_i}. \tag{51}$$

Asymptotic Normality of Estimators

It follows from (51), (50), and (49) that

$$\theta_n^*(a_{n\bullet}, b_{n\bullet}) - \theta = \frac{\sum_{i=1}^n u_{ni}\tilde{\varepsilon}_i}{\beta + \sum_{i=1}^n b_{ni}\tilde{\varepsilon}_i/B_n}, \quad \text{where} \quad u_{ni} := \frac{a_{ni}}{A_n} - \theta\frac{b_{ni}}{B_n}.$$

So, we are ready to prove the asymptotic normality of θ^*.

Theorem 8 (See [5]) *Suppose that conditions* (44)–(47) *hold. Furthermore assume that*

$$\max_{i \le n} u_{ni}^2 (1 + x_i)^2 \Big/ \sum_{i=1}^{n} (1 + x_i)^2 u_{ni}^2 \to 0,$$

$$Q_n(b_{n\bullet}) := \sum_{i=1}^{n} b_{ni}^2 (1 + x_i)^2 \Big/ B_n^2 \to 0.$$

Then $\theta_n^*(a_{n\bullet}, b_{n\bullet})$ is asymptotically normal:

$$\frac{\theta_n^*(a_{n\bullet}, b_{n\bullet}) - \theta}{\tilde{\sigma} d_n(a_{n\bullet}, b_{n\bullet})} \Rightarrow \mathcal{N}(0, 1), \quad \textit{where} \quad d_n^2(a_{n\bullet}, b_{n\bullet}) := \sum_{i=1}^{n} u_{ni}^2 (1 + x_i\theta)^2.$$

Optimization of Estimators

For all $\alpha > 0$ introduce notations

$$C_{kn}(\alpha) := \sum_{i=1}^{n} \frac{x_i^k}{(1 + x_i\alpha)^2}, \quad k = 0, 1, 2.$$

$$\bar{a}_{ni}(\alpha) := \frac{C_0(\alpha)x_i - C_1(\alpha)}{(1 + x_i\alpha)^2}, \quad \bar{b}_{ni}(\alpha) := \frac{C_2(\alpha) - C_1(\alpha)x_i}{(1 + x_i\alpha)^2}, \quad i = 1, \ldots, n. \tag{52}$$

It is not difficult to verify that for all $\alpha > 0$ numbers $\{a_{ni} = \bar{a}_{ni}(\alpha)\}$ and $\{b_{ni} = \bar{b}_{ni}(\alpha)\}$ satisfy assumptions (46) and (47) with $A_n = B_n = \Delta_n(\alpha)$ where

$$\Delta_n(\alpha) := C_{0n}(\alpha)C_{2n}(\alpha) - C_{1n}^2(\alpha) = C_{0n}(\alpha) \sum_{i=1}^{n} \frac{\big(x_i - C_{1n}(\alpha)/C_{0n}(\alpha)\big)^2}{(1 + x_i\alpha)^2}.$$

And $\Delta_n(\alpha) > 0$ for all $n \ge n_0 = \min\{n : x_n \ne x_1\}$.

Theorem 9 (See [5]) *If θ is the true value of the main unknown parameter and $n \ge n_0$ then for all numbers $\{a_{ni}\}$ and $\{b_{ni}\}$ satisfying assumptions* (46) *and* (47)*, the following inequalities hold:*

$$d_n^2(a_{n\bullet}, b_{n\bullet}) \ge d_{n,opt}^2(\theta) := n/\Delta_n(\theta) = d_n^2(\bar{a}_{n\bullet}(\theta), \bar{b}_{n\bullet}(\theta)),$$
$$Q_n(b_{n\bullet}) \ge Q_{n,opt}(\theta) := C_{2n}(\theta)/\Delta_n(\theta) = Q_n(\bar{b}_{n\bullet}(\theta)).$$

Partial Cases

For an arbitrary $\alpha > 0$ consider an estimator $\theta_{n,\alpha}^* = \theta_n^*(\bar{a}_{n\bullet}(\alpha), \bar{b}_{n\bullet}(\alpha))$. Note $\theta_{n,\alpha}^*$ is the estimator which we obtain once we set $\{a_{ni} = \bar{a}_{ni}(\alpha)\}$ and $\{b_{ni} = \bar{b}_{ni}(\alpha)\}$ in the definition (52). It follows from Theorem 9 that this estimator has the minimal asymptotic variance in case when $\theta = \alpha$. Due to this reason, we may recommend to use it at least in the case when we know that α is in some sense close to the unknown θ.

In particular, we may recommend to use the estimator

$$\theta^*_{n,0} := \frac{n \sum_{i=1}^n x_i Z_i - \sum_{i=1}^n x_i \sum_{i=1}^n Z_i}{\sum_{i=1}^n x_i^2 \sum_{i=1}^n Z_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i Z_i},$$

if we suppose that θ is small in some sense. And for a large θ we may apply

$$\theta^*_{n,\infty} := \frac{\sum_{i=1}^n 1/x_i^2 \sum_{i=1}^n Z_i/x_i - \sum_{i=1}^n 1/x_i \sum_{i=1}^n Z_i/x_i^2}{n \sum_{i=1}^n Z_i/x_i^2 - \sum_{i=1}^n 1/x_i \sum_{i=1}^n Z_i/x_i}.$$

Remark 6 By Theorem 9 the random variable $\theta^*_{n,\theta}$ minimizes the asymptotic variance in the class of random variables defined in (51) for all θ and at the same time it makes the quantity $Q_n(b_{n\bullet})$ in Theorem 8, as small as possible. Unfortunately, we cannot use this variable $\theta^*_{n,\theta}$ as an estimator since it depends on the unknown parameter θ.

Possible Generalizations

Remark 7 If $\mathbf{E}\delta_1 = 0$ then we may try to improve the estimator θ^*_n using ideas from Sect. 8 with $g_i(t) = \log(1 + x_i t)$ and $w_i(t) = 1$.

Remark 8 Of course, all these results may be extended to the case when

$$Y_i = \log(x_{0i} + x_{1i}\theta_1 + \cdots + x_{mi}\theta_m) + \varepsilon_i,$$

where $(\theta_1, \ldots, \theta_m)$ is an m-dimensional unknown parameter.

6 General Power Regression

Supposed that we have observations which may be represented in the following form

$$Y_i = (z_i + s_i\theta)^r + \varepsilon_i = (z_i + s_i\theta)^r(1 + \delta_i) > 0, \qquad i = 1, 2, \ldots, \tag{53}$$

where $\{z_i, s_i > 0\}$ and $r \neq 0$ are known numbers. We assume that unobservable random variables $\delta_1, \delta_2, \ldots$ are i.i.d. with

$$0 < \text{var } (1 + \delta_1)^{1/r} < \infty \quad \text{and} \quad n_1 := \min\{n : s_n/z_n \neq s_1/z_1\} < \infty.$$

Our aim is to construct a simple explicit estimator for the unknown parameter $\theta > 0$ using the first $n \geq n_1$ observations from (53).

Construction of Explicit Estimators

First of all, we rewrite (53) in the following way:

$$Y_i^{1/r} = (z_i + s_i\theta)(1+\delta_i)^{1/r}, \qquad i = 1, 2, \ldots. \tag{54}$$

Introduce notations:

$$Z_i := Y_i^{1/r}/z_i, \quad x_i := s_i/z_i, \quad \beta := \mathbf{E}(1+\delta_i)^{1/r}, \quad \tilde{\delta}_i := (1+\delta_i)^{1/r}/\beta - 1. \tag{55}$$

Here we have defined a new parameter β. Notations from (55) allow us to rewrite (54) in the following form:

$$Z_i = \beta(1 + x_i\theta)(1+\tilde{\delta}_i) = \beta + x_i\beta\theta + \tilde{\varepsilon}_i, \qquad i = 1, 2, \ldots.$$

But this formula coinsides with (45).

Let us stress that under notations from (55), Eqs. (53) and (54) are equivalent to (45). Moreover, it is easy to see that conditions (44) hold again for $\{\tilde{\delta}_i\}$ introduced in (55).

Later on in this section we will use estimators $\theta_n^* = \theta_n^*(a_{n\bullet}, b_{n\bullet})$ introduced in (50) with any numbers $\{a_{ni}\}$ and $\{b_{ni}\}$ satisfying (46) and (47). But we assume that in this section all $\{Z_i\}$ and $\{x_i\}$ are from (55).

Main Assertions

Theorem 10 *All assertions of Theorems 8 and 9 are true with $\{Z_i\}$ and $\{x_i\}$ defined in (55).*

Remark 9 Careful study of the procedures of the estimators' construction in Sects. 5 and 6 shows quite easily that in Sect. 5 we actually constructed the estimator θ_n^* for the parameter θ from the simplified Eq. (45), which is equivalent to the original Eq. (42). Similarly, in Sect. 6 we estimated the parameter from the simplified Eq. (45), which is now turned out to be equivalent to the considered in this section Eq. (53).

Note that the properties of the estimators θ_n^* in Sects. 5 and 6 coincide since we only use the same assumptions (44) and (45). This observation immediately leads to Theorem 10.

Furthermore we note that all derivations in Sect. 5 from Theorem 10 remain true for the estimators considered in the present Sect. 6.

Possible Generalizations

Remark 10 If $\mathbf{E}\delta_1 = 0$ then we may try to improve the estimator θ_n^* using the ideas from Sect. 8 with functions $g_i(t) = (z_i + s_i t)^r$ and $w_i(t) = 1/g_i(t)$.

Remark 11 Of course, all these results may be extended to the case when

$$Y_i = (x_{0i} + x_{1i}\theta_1 + \cdots + x_{mi}\theta_m)^r(1+\delta_i) > 0, \qquad i = 1, 2, \ldots,$$

where $(\theta_1, \ldots, \theta_m)$ is an m-dimensional unknown parameter.

7 Exponential Regression

Suppose that we have observations of the form

$$Y_i = e^{x_i\theta} + \varepsilon_i = e^{x_i\theta}(1+\delta_i) > 0, \qquad i = 1, 2, \ldots, \tag{56}$$

where $\{x_i > 0\}$ are known numbers. We assume that unobservable random variables $\delta_1, \delta_2, \ldots$ are i.i.d. with

$$0 < \tilde{\sigma}^2 := \text{var} \ \log(1+\delta_1) < \infty \quad \text{and} \quad n_0 := \min\{n : x_n \neq x_1\} < \infty.$$

Our aim is to construct a simple explicit estimator for the main unknown parameter $\theta > 0$ using the first $n \geq n_0$ observations from (56).

Construction of Explicit Estimators

First of all, we rewrite (56) in the following way:

$$Z_i := \log Y_i = x_i\theta + \beta + \tilde{\delta}_i, \qquad i = 1, 2, \ldots, \tag{57}$$

where we use notations:

$$\beta := \mathbf{E}\log(1+\delta_1), \qquad \tilde{\delta}_i := \log(1+\delta_i) - \beta. \tag{58}$$

It is easy to see that random variables $\{\tilde{\delta}_i\}$ again satisfy conditions (44).

Let us stress that the Eq. (57) is the standard equation of the linear regression with two unknown parameters. So, we may take the following estimator:

$$\theta_n^* := \sum_{i=1}^n Z_i(x_i - \bar{x}_n) \Big/ \sum_{i=1}^n (x_i - \bar{x}_n)^2, \tag{59}$$

which coincides with the famous Least Square Estimator for the parameter θ in Eq. (57), and due to this reason it is optimal in some sense. In (59) and later on we use the following notations:

$$\bar{x}_n := \sum_{i=1}^n x_i \Big/ n, \qquad s_n^2 := \sum_{i=1}^n (x_i - \bar{x}_n)^2.$$

After that we may introduce the following more complicated estimator:

$$\theta_n^{**} := \theta_n^* + \sum_{i=1}^n x_i \left(Y_i e^{-x_i \theta_n^*} - 1\right) \Big/ \sum_{i=1}^n x_i^2. \tag{60}$$

Asymptotic Normality of Estimators

Theorem 11 *Suppose that conditions* (44) *hold and*

$$s_n^2 \to \infty, \qquad \max_{i \le n} (x_i - \bar{x}_n)^2 \big/ s_n^2 \to 0. \tag{61}$$

Then θ_n^* is unbiased and asymptotically normal:

$$(\theta_n^* - \theta)/d_n \Rightarrow \mathcal{N}(0, 1), \quad \textit{where} \quad d_n^2 := \tilde{\sigma}^2 \big/ s_n^2 \to 0. \tag{62}$$

Theorem 12 *Suppose that assumptions* (44) *hold and*

$$s_n^2 \to \infty, \qquad \left(\max_{i \le n} x_i^2 + \sqrt{n} \bar{x}_n^2\right) \big/ s_n^2 \to 0, \tag{63}$$

$$\{\delta_i\} \textit{ are i.i.d. with } \mathbf{E}\delta_1 = 0 \quad \textit{and} \quad 0 < \sigma^2 := \text{var } \delta_1 < \infty. \tag{64}$$

Then (62) *takes place and θ_n^{**} is also asymptotically normal:*

$$\frac{\theta_n^{**} - \theta}{D_n} \Rightarrow \mathcal{N}(0, 1), \quad \textit{where} \quad D_n^2 := \frac{\sigma^2}{\sum_{i=1}^n x_i^2} = \frac{\sigma^2}{s_n^2 + n\bar{x}_n^2} \to 0. \tag{65}$$

Remark 12 Both assumptions in (61) follow from conditions (63) of Theorem 12 since

$$\sum_{i=1}^n x_i^2 = s_n^2 + n\bar{x}_n^2. \tag{66}$$

In particular,

$$\text{if} \quad n\bar{x}_n^2/s_n^2 \to \infty \quad \text{then} \quad D_n^2/d_n^2 \to 0. \tag{67}$$

So, in this case the estimator θ_n^{**} is sharper than θ_n^* for sufficiently large n.

Moreover,

$$\text{if} \quad \sigma^2 \le \tilde{\sigma}^2 \quad \text{then} \quad D_n^2 \le d_n^2 \quad \forall n \ge n_0. \tag{68}$$

On the other hand, if

$$\tilde{\sigma}^2 \le \sigma^2 \quad \text{and} \quad n\bar{x}_n^2 \big/ s_n^2 \to 0, \tag{69}$$

then the estimator θ_n^* is sharper than θ_n^{**} for sufficiently large n.

Examples

Here we present examples when all the situations (67)–(69) take place.

Example 4 Let for all $k = 1, 2, \ldots$

$$x_k = 1 - 1/k^\alpha, \quad \text{where} \quad 0 < \alpha < 1/4. \tag{70}$$

Then (67) holds and

$$d_n^2 = (1-2\alpha)(1-\alpha)^2 n^{2\alpha-1}/\alpha^2 \to 0, \qquad D_n^2 \sim \sigma^2/n = o(d_n^2). \tag{71}$$

Example 5 Let $\tilde{\delta}_1$ have a normal distribution with zero mean. Then the first assumption in (69) is true since

$$\sigma^2 = \text{var } \delta_1 = e^{\tilde{\sigma}^2} - 1 > \tilde{\sigma}^2 = \text{var } \tilde{\delta}_1 > 0. \tag{72}$$

Example 6 Let

$$\mathbf{P}(\tilde{\delta}_1 = \tilde{\sigma}) = 1/2 = \mathbf{P}(\tilde{\delta}_1 = -\tilde{\sigma}). \tag{73}$$

Using hyperbolic functions we obtain that

$$0 < \sigma^2 = \text{var}\, \delta_1 = \tanh^2 \tilde{\sigma} < \tilde{\sigma}^2 = \text{var}\, \tilde{\delta}_1. \tag{74}$$

So, (68) takes place.

Assertions of Theorems 11 and 12 will be proved in Sect. 11 together with the facts from Examples 4–6.

8 General Remarks About One-Dimensional Estimators

If we constructed a "sufficiently good" initial estimator θ_n^* and want to find an estimator with property (15) then, first of all, we may consider a class of explicit estimators defined by the formula:

$$\theta_n^{**} = \theta_n^* + \frac{\sum_{i=1}^n h_i(\theta_n^*)\varepsilon_i^*}{\sum_{i=1}^n h_i(\theta_n^*) g_i'(\theta_n^*) + \sum_{i=1}^n \mu_i(\theta_n^*)\varepsilon_i^*} \quad \text{with} \quad \varepsilon_i^* = Y_i - g_i(\theta_n^*) \tag{75}$$

for some sufficiently smooth functions $\{h_i(\cdot)\}$ and $\{\mu_i(\cdot)\}$.

In [6, 7] it is easy to find many arguments to use such estimators with

$$h_i(t) = w_i(t) g_i'(t) \quad \text{if} \quad \text{var } \varepsilon_i = \sigma^2/w_i(\theta), \qquad i = 1, 2, \ldots. \tag{76}$$

For example, using Taylor formula it is possible to obtain from (75) that under appropriate conditions

$$\theta_n^{**} - \theta \approx \frac{\sum_{i=1}^n h_i(\theta)\varepsilon_i}{\sum_{i=1}^n h_i(\theta_n) g_i'(\theta)}. \tag{77}$$

So, under some assumptions θ_n^{**} is asymptotically normal:

$$(\theta_n^{**} - \theta_n^*)/D_n \Rightarrow \mathcal{N}(0, 1) \quad as\ n \to \infty, \tag{78}$$

where the asymptotic variance D_n^2 has the following form

$$D_n^2 = \frac{\sum_{i=1}^n h_i^2(\theta) \operatorname{var}\ \varepsilon_i}{\left(\sum h_i(\theta) g_i'(\theta)\right)^2} \geq \frac{1}{\sum_{i=1}^n (g_i'(\theta))^2 \big/ \operatorname{var}\ \varepsilon_i}. \tag{79}$$

Thus, under assumptions (76) we have $D_n = D_{n,o}$ with $D_{n,o}$ defined in (14).

So, if conditions (13) are satisfied and $h_i(t) = w_i(t) g_i'(t)$ are optimal, then we have the equality in (79). If $\{w_i(\cdot)\}$ are known only approximately, or $\{g_i'(\cdot)\}$ are not sufficiently smooth, then we may try to use smooth approximations $\{h_i(t)\}$ of such optimal functions $\{w_i(t) g_i'(t)\}$.

Let us also note another argument in favor of the estimators from (75). Denote by $\check{\theta}_n$ the so called "quasy likelihood estimator" (see [6] or [7]) which we define as the solution of the following equation:

$$\sum_{i=1}^n h_i(\check{\theta}_n)(Y_i - g_i(\check{\theta}_n)) = 0. \tag{80}$$

Substituting (1) into (80) and using Taylor formula yield easily

$$0 = \sum_{i=1}^n h_i(\check{\theta}_n)(\varepsilon_i + g_i(\theta) - g_i(\check{\theta}_n)) \approx \sum_{i=1}^n h_i(\theta)\varepsilon_i + (\theta - \check{\theta}_n) \sum_{i=1}^n h_i(\theta) g_i'(\theta)$$

under appropriate conditions. So, relations (77)–(79) remain true with $\check{\theta}_n$ instead of θ_n^{**}. And it follows from (79) that functions $h_i(t) = w_i(t) g_i'(t)$ from and (76) are again optimal.

In this case the estimator θ_n^{**} may be treated as a Newton–Raphson approximation of $\check{\theta}_n$. Note, that in partial cases interesting investigations of distances between estimators $\check{\theta}_n$ and θ_n^{**} may be found in [8, 9].

Remark 13 It follows from the constructions in Sects. 2–7 that explicit asymptotically normal estimators for the parameter of Eq. (1) may exist only for a limited number of "lucky" functions $\{g_i(\cdot)\}$ when some special tricks with this Eq. (1) can be performed.

Moreover, in case of Eqs. (9)–(12) special assumptions about $\{\operatorname{var}\ \varepsilon_i\}$ must take place together with the property that $\{\delta_i := \varepsilon_i / \sqrt{\operatorname{var}\ \varepsilon_i}\}$ are i.i.d.

Remark 14 The presence of significant additional conditions on the distributions of the variables $\{\varepsilon_i\}$ makes possible the existence of estimators which are better than θ_n^{**} with the condition (15). For this reason, in some of the problems (9)–(12) the estimator θ_n^{**} with the condition (15) may be worse than the estimator θ_n^*. The possibility of this at the first glance unexpected situation follows from Example 5 and Remark 12. By the last reasoning in Sects. 4–7, we avoid calling the estimators θ_n^{**} as the improvements of the estimators θ_n^*.

Remark 15 Estimator (60), which was introduced and investigated in Sect. 7, is a partial case of the estimators from (75) for the Eq. (12). The author in collaboration with his students is planning to perform similar investigation for Eqs. (8)–(11) too.

Note that the estimator introduced in (25) is not a partial case of the estimators from (75) for the Eq. (18). Nevertheless, the optimal version of this estimator satisfies the equality $d_{n,opt}^2 = D_{n,o}^2$ as it follows from (26) and (14).

9 Multidimensional Case

Suppose that we have observations

$$Y_i = \frac{\alpha_i(\theta)}{\beta_i(\theta)} + \varepsilon_i, \qquad i = 1, \dots, n, \tag{81}$$

where $\theta = (\theta_1, \dots, \theta_m)$ is the unknown parameter which we want to estimate and

$$\alpha_i(\theta) \equiv a_{0i} + \sum_{j=1}^{m} a_{ji}\theta_j, \qquad \beta_i(\theta) \equiv 1 + \sum_{j=1}^{m} b_{ji}\theta_j$$

with $\theta_j > 0$ and $b_{ji} \geq 0$ for all j and i.

Construction of Explicit Estimators

Rewrite (81) as

$$Y_i + \sum_{j=1}^{m} b_{ji} Y_i \theta_j = a_{0i} + \sum_{j=1}^{m} a_{ji}\theta_j + \beta_i(\theta)\varepsilon_i. \tag{82}$$

Introduce notations

$$X_{ji} = a_{ji} - b_{ji}Y_i, \quad Z_i = Y_i - a_{0i}, \quad \tilde{\varepsilon}_i = \beta_i(\theta)\varepsilon_i.$$

So, (81) is equivalent to (82) and is equivalent to

$$Z_i = \sum_{j=1}^{m} X_{ji}\theta_j + \tilde{\varepsilon}_i, \qquad i = 1, \ldots, n. \tag{83}$$

For any constants $\{c_{ki}\}$ we have from (83) that

$$\sum_{i=1}^{n} c_{ki} Z_i - \sum_{j=1}^{m} \theta_j \sum_{i=1}^{n} c_{ki} X_{ji} = \sum_{i=1}^{n} c_{ki}\tilde{\varepsilon}_i, \tag{84}$$

where $k = 1, \ldots, m$ and $i = 1, \ldots, n$. Now, define the estimator $\theta^* = (\theta_1^*, \ldots, \theta_m^*)$ as the solution of m linear equations:

$$\sum_{j=1}^{m} \theta_j^* \sum_{i=1}^{n} c_{ki} X_{ji} = \sum_{i=1}^{n} c_{ki} Z_i, \qquad k = 1, \ldots, m. \tag{85}$$

After that we may choose appropriate functions $\{\gamma_{ki}(\cdot)\}$ and introduce improved estimators $\theta^{**} = (\theta_1^{**}, \ldots, \theta_m^{**})$ as the solution of the following m linear equations:

$$\sum_{j=1}^{m} \theta_j^{**} \sum_{i=1}^{n} \gamma_{ki}(\theta^*) X_{ji} = \sum_{i=1}^{n} \gamma_{ki}(\theta^*) Z_i, \qquad k = 1, \ldots, m. \tag{86}$$

Remark 16 Thus, in the case of the m-dimensional unknown parameter we call the solutions of m linear equations whose coefficients are previously found explicit statistics, which do not depend on any unknown parameters or statistics, as the explicit estimators.

Asymptotic Normality and Optimization of Estimators

It follows immediately from (84) and (85) that

$$\sum_{j=1}^{m} (\theta_j^* - \theta_j) \sum_{i=1}^{n} c_{ki} X_{ji} = \sum_{i=1}^{n} c_{ki}\tilde{\varepsilon}_i, \quad k = 1, \ldots, m.$$

Using (86) we may obtain similar representations for the differences $\{\theta_j^{**} - \theta_j\}$. Thus, we are ready to prove the asymptotic normality of estimators $\theta^* = (\theta_1^*, \ldots, \theta_m^*)$ and $\theta^{**} = (\theta_1^{**}, \ldots, \theta_m^{**})$. Using assumption (13) introduce notation

$$\gamma_{ji,opt}(\theta) := \big(a_{ji} - b_{ji}\alpha_i(\theta)/\beta_i(\theta)\big) w_i(\theta). \tag{87}$$

Theorem 13 *Let* (81) *and* (13) *hold. Then under appropriate additional assumptions the estimators* $\theta^* = (\theta_1^*, \ldots, \theta_m^*)$ *and* $\theta^{**} = (\theta_1^{**}, \ldots, \theta_m^{**})$ *are asymptotically normal with some asymptotic covariance matrices* $\mathbf{B}(\{c_{ji}\})$ *and* $\mathbf{B}(\{\gamma_{ji}(\theta)\})$ *correspondingly. Moreover, the matrices*

$$\mathbf{B}(\{c_{ji}\}) - \mathbf{B}(\{\gamma_{ji,opt}(\theta)\}) \quad \textit{and} \quad \mathbf{B}(\{\gamma_{ji}(\theta)\}) - \mathbf{B}(\{\gamma_{ji,opt}(\theta)\})$$

are nonnegatively definite.

Note, that numbers $\{c_{ki}\}$ and functions $\{\gamma_{ki}(\cdot)\}$ used in the definitions (85) and (86) may depend on the number n of observations whereas the optimal functions from (87) which we recommend to use in (86) does not depend on n.

The proof of Theorem 13 together with detailed descriptions of the associated asymptotic covariance matrices may be found in [10].

A partial Case

The following example is taken from [7, p. 77].

Example 7 Suppose that we have independent observations

$$Y_i = \frac{V\theta_2(x_{2i} - x_{3i}/1.632)}{1 + x_{1i}\theta_1 + x_{2i}\theta_2 + x_{3i}\theta_3} + \varepsilon_i, \qquad i = 1, \dots, n, \tag{88}$$

where $\{x_{1i},\ x_{2i},\ x_{3i} > 0\}$ are given numbers. Our aim is to estimate the unknown parameters $\theta_1, \theta_2, \theta_3, V > 0$.

Introduce new variables

$$\theta_4 := V\theta_2 \quad \text{and} \quad x_{4i} = x_{2i} - x_{3i}/1.632, \quad i = 1, \dots, n.$$

So, Eq. (88) turn into a particular case of Eq. (81) and we may find estimators $\theta_1^*, \theta_2^*, \theta_3^*, \theta_4^*$ using a version of the system (85) with four linear equations. After that we need only to define

$$V^* := \theta_4^* \theta_2^* \quad \text{since} \quad \theta_4 = V\theta_2.$$

The most known partial case of Eq. (81) is the famous Michaelis–Menten equation, which will be considered in the following section.

10 Michaelis–Menten Equation

For some known $s_i > 0$ let

$$Y_i = \frac{Vs_i}{K + s_i} + \varepsilon_i = \frac{V}{1 + K/s_i} + \varepsilon_i, \quad i = 1, 2, \dots. \tag{89}$$

We want to estimate unknown parameters $K > 0$ and $V > 0$.

Construction of Explicit Estimators

Rewrite (89) in the following way:

$$KY_i/s_i + Y_i = V + \tilde{\varepsilon}_i \quad \text{with} \quad \tilde{\varepsilon}_i = (1 + Kx_i)\varepsilon_i. \tag{90}$$

For any constants $\{a_i\}$ and $\{b_i\}$ such that

$$\sum_{i=1}^{n} a_i = 0 \qquad \text{and} \qquad \sum_{i=1}^{n} b_i = 1$$

we have from (90)

$$K\sum_{i=1}^{n} a_i Y_i/s_i + \sum_{i=1}^{n} a_i Y_i = 0 + \sum_{i=1}^{n} a_i \tilde{\varepsilon}_i,$$

$$K\sum_{i=1}^{n} b_i Y_i/s_i + \sum_{i=1}^{n} b_i Y_i = V \cdot 1 + \sum_{i=1}^{n} b_i \tilde{\varepsilon}_i.$$

Ignoring the sum of the approximation errors in the equalities above, it is natural to define K^* and V^* as solutions of equations:

$$K^*\sum_{i=1}^{n} a_i Y_i/s_i + \sum_{i=1}^{n} a_i Y_i = 0, \qquad K^*\sum_{i=1}^{n} b_i Y_i/s_i + \sum_{i=1}^{n} b_i Y_i = V^*.$$

So, we have

$$K^* = -\sum_{i=1}^{n} a_i Y_i \Big/ \sum_{i=1}^{n} a_i Y_i/s_i, \qquad V^* = \sum_{i=1}^{n} b_i (1 + K^*/s_i) Y_i.$$

In particular, if $b_i = 1/n$ and $a_i = s_i - \bar{s}_n$ with $\bar{s}_n = n^{-1}\sum_{i=1}^{n} s_i$ then

$$K^* = \sum_{i=1}^{n} (\bar{s}_n - s_i) Y_i \Big/ \sum_{i=1}^{n} (1 - \bar{s}_n/s_i) Y_i, \qquad V^* = n^{-1}\sum_{i=1}^{n} (Y_i + K^* Y_i/s_i).$$

Construction of Optimal Explicit Improved Estimators

It is evident that Eq. (89) is a partial case of (88). Hence we may apply Theorem 13 and find optimal improved estimators K^{**} and V^{**}.

Suppose that assumption (13) holds then rewrite it in the following form:

$$\mathbf{E}\varepsilon_i = 0 \quad \text{and} \quad 0 < \text{var}\ \varepsilon_i = \sigma^2/w_i(K, V) < \infty, \quad i = 1, 2, \ldots.$$

Introduce

$$\gamma_i^* := w_i(K^*, V^*)/(1 + K^*/s_i)^3, \quad i = 1, 2, \ldots.$$

Define now estimators K^{**} and V^{**} as solutions of the following system with two linear equations:

$$K^{**}\sum_{i=1}^{n}\gamma_i^* Y_i/s_i + \sum_{i=1}^{n}\gamma_i^* Y_i = V^{**}\sum_{i=1}^{n}\gamma_i^*, \tag{91}$$

$$K^{**}\sum_{i=1}^{n}\gamma_i^* Y_i/s_i^2 + \sum_{i=1}^{n}\gamma_i^* Y_i/s_i = V^{**}\sum_{i=1}^{n}\gamma_i^*/s_i. \tag{92}$$

It was shown in [11] that two-dimensional estimator (K^{**}, V^{**}) is asymptotically optimal in the sense of Theorem 13.

Remark 17 If for some known $\{w_{0i} > 0\}$

$$\text{var}\ \varepsilon_i = \frac{\sigma^2}{w_{0i}(1 + K/s_i)^3}, \qquad i = 1, 2, \ldots,$$

then $\gamma_i^* = w_{0i}$ for all i and we may find optimal estimators (K^{**}, V^{**}) as solutions of Eqs. (91) and (92).

Note that it is the only case where one does not need to search for initial estimators.

Remark 18 Since 1913 (see [12]) Michaelis–Menten equation is popular in biochemistry where it describes rates of enzymatic reactions. The most interesting cases for chemistry are var $\varepsilon_i = \sigma^2 \big/ (1 + K/s_i)^2$ and var $\varepsilon_i = \sigma^2$.

On Explicit Estimators of Johansen and Lumry

The author's interest to search for explicit estimators in Michaelis–Menten equation is motivated by the following explicit estimators

$$\tilde{V} = \frac{\sum_{i=1}^n Y_i^2/s_i^2 \sum_{i=1}^n Y_i^2 - \left(\sum_{i=1}^n Y_i^2/s_i\right)^2}{\sum_{i=1}^n Y_i^2/s_i^2 \sum_{i=1}^n Y_i - \sum_{i=1}^n Y_i^2/s_i \sum_{i=1}^n Y_i/s_i},$$

$$\tilde{K} = \frac{\sum_{i=1}^n Y_i^2 \sum_{i=1}^n Y_i/s_i - \sum_{i=1}^n Y_i^2/s_i \sum_{i=1}^n Y_i}{\sum_{i=1}^n Y_i^2/s_i^2 \sum_{i=1}^n Y_i - \sum_{i=1}^n Y_i^2/s_i \sum_{i=1}^n Y_i/s_i}$$

which may be found, for example, in Chap. 10 of [13]. There these estimators were derived from some heuristic considerations if the following assumptions hold

$$\mathbf{E}\varepsilon_i = 0 \quad \text{and} \quad 0 < \text{var}\ \varepsilon_i = \sigma^2(1 + K/s_i)^2 < \infty, \quad i = 1, 2, \ldots. \tag{93}$$

However, a more detailed theoretical investigation shows a series of shortcomings of these estimators.

Theorem 14 *Let the values* $\{s_i\}$ *be i.i.d. random variables with a nondegenerate common distribution. Suppose that independent random variables* $\{\varepsilon_i\}$ *have normal distributions and* (93) *holds. Additionally, assume that sequences* $\{s_i\}$ *and* $\{\varepsilon_i\}$ *are independent. Then*

$$\tilde{V} \xrightarrow{p} V + \sigma^2/V \neq V.$$

So, $\tilde{V}$ is inconsistent.

This statement is proven in [14].

11 Proofs

We need only to prove several assertions from Sect. 7.

Proof of Theorem 11

Substituting (57) into (59) yields

$$\delta^* := \theta_n^* - \theta = \sum_{i=1}^{n} (x_i - \bar{x}_n)\tilde{\delta}_i / s_n^2.$$

So, from (57) we obtain:

$$\frac{\theta_n^* - \theta}{d_n} = \sum_{i=1}^{n} \frac{c_{ni}\tilde{\delta}_i}{\tilde{\sigma}} \quad \text{with} \quad c_{ni} = \frac{x_i - \bar{x}_n}{s_n}. \tag{94}$$

The following fact is well known.

Lemma 1 *Let* $\eta_1, \eta_2, \ldots$ *be i.i.d. random variables with* $\mathbf{E}\eta_1 = 0$ *and* var $\eta_1 = 1$. *Assume that for all n we are given the real numbers* $\{c_{ni}\}$ *such that*

$$\sum_{i=1}^{n} c_{ni}^2 = 1 \quad \text{and} \quad \max_{i \le n} c_{ni}^2 \to 0. \tag{95}$$

Then $\sum_{i=1}^{n} c_{ni}\eta_i \Rightarrow \mathcal{N}(0, 1)$,

Lemma 1 with $\eta_1 = \tilde{\delta}_i/\tilde{\sigma}$ and the representation (94) yield the convergence (62) since the condition (95) follows from (61).

Representation for θ_n^{**}

Below in this section we suppose that all assumptions of Theorem 12 are fulfilled. Let

$$z_n^2 := \sum_{i=1}^{n} x_i^2, \quad f_i(t) = e^{-x_i t} - 1 + x_i t. \tag{96}$$

It follows from (63), (66) and Lemma 1 with $\eta_i = \delta_i/\sigma$ and $c_{ni} = x_i/z_n$ that

$$\rho_0 := \sum_{i=1}^{n} x_i\delta_i/(\sigma z_n) \Rightarrow \mathcal{N}(0,1). \tag{97}$$

Now, by substituting (56) into (60) we obtain:

$$\theta_n^{**} - \theta = \delta^* + \sum_{i=1}^{n} x_i(e^{-x_i\delta^*}(1+\delta_i) - 1)/z_n^2. \tag{98}$$

Introduce notations:

$$\rho_1 := \sum_{i=1}^{n} x_i f_i(\delta^*), \quad \rho_2 := \sum_{i=1}^{n} x_i f_i(\delta^*)\delta_i, \quad \rho_3 := \sum_{i=1}^{n} x_i^2\delta_i. \tag{99}$$

Now, we may rewrite (98) in the following way:

$$\rho^{**} := \frac{\theta_n^{**} - \theta}{D_n} = \rho_0 + \frac{\rho_1 + \rho_2 - \delta^*\rho_3}{\sigma z_n}, \tag{100}$$

since $D_n = \sigma/z_n$ by (65).

Auxiliary Lemmas

Introduce notations

$$\lambda_n := \max_{i\le n}|x_i|, \quad \rho_4 := \sum_{i=1}^{n}|x_i|^3, \quad \rho_5 := \sum_{i=1}^{n}|x_i|^3|\delta_i|.$$

Lemma 2 *For all $n \ge n_0$*

$$|\rho_1| \le |\delta^*|^2 e^{\lambda_n|\delta^*|}\rho_4, \quad |\rho_2| \le |\delta^*|^2 e^{\lambda_n|\delta^*|}\rho_5. \tag{101}$$

In addition

$$\rho_4 \le 8\lambda_n s_n z_n + 4\sqrt{n} z_n \bar{x}_n^2, \tag{102}$$

$$\mathbf{E}\rho_5 \le \sigma\rho_4, \quad \mathbf{E}|\rho_3| \le \sigma\lambda_n z_n, \quad \mathbf{E}|\delta_n^*| \le \sigma/s_n. \tag{103}$$

Proof Using Taylor formula we easily obtain from the definition (96) that for all $i \le n$

$$|f_i(\delta^*)| \le |x_i\delta^*|^2 e^{|x_i\delta^*|}/2 \le x_i^2|\delta^*|^2 e^{\lambda_n|\delta^*|}.$$

This fact and definitions (99) give us inequalities (101).

It is easy to see that

$$x_i = (x_i - \bar{x}_n) + \bar{x}_n \quad \text{and} \quad |x_i|^3 \le 4|x_i - \bar{x}_n|^3 + 4|\bar{x}_n|^3.$$

Hence

$$\rho_4 \le 4\sum_{i=1}^{n} |x_i - \bar{x}_n|^3 + 4n|\bar{x}_n|^3 \le 8\lambda_n s_n^2 + 4\sqrt{n} z_n \bar{x}_n^2. \tag{104}$$

Here we used that $|x_i| \le \lambda_n$ for $i \le n$ and that $|\bar{x}_n| \le \lambda_n$, $n\bar{x}_n^2 \le z_n^2$. We obtain (102) from (104) since $s_n \le z_n$ by (66).

Estimates (103) follow from the corresponding definitions and the following facts:

$$\mathbf{E}\delta_i^2 = 1, \quad \mathbf{E}\rho_3^2 = \sigma^2 \sum_{i=1}^{n} x_i^4 \le \sigma^2 \lambda_n^2 z_n^2, \quad \mathbf{E}|\delta_n^*|^2 = \sigma^2/s_n^2.$$

Lemma 3 *If random variables $\{\zeta_n\}$ are such that $\mathbf{E}|\zeta_n| < \infty$ for all sufficiently large n then*

$$\zeta_n = O_p(\mathbf{E}|\zeta_n|).$$

This fact follows from the Chebyshev's inequality with the first moment.

Proof of Theorem 12

It follows from (103) and Lemma 3 that

$$\rho_5 = O_p(\rho_4), \quad \rho_3 = O_p(\lambda_n z_n), \quad \delta_n^* = O_p(1/s_n). \tag{105}$$

Since $\lambda_n/s_n \to 0$ by condition (63), we have that $\lambda_n|\delta^*| = O_p(\lambda_n/s_n) = o_p(1)$. This fact together with (103) and (105) yield

$$\rho_1 = O_p(\rho_4/s_n^2), \quad \rho_2 = O_p(\rho_4/s_n^2), \quad \delta^*\rho_3 = O_p(\lambda_n z_n/s_n).$$

Now, from these relationships and (100) we obtain:

$$z_n(\rho^{**} - \rho_0) = O_p(\rho_4/s_n^2 + \rho_4/s_n^2 + \lambda_n z_n/s_n).$$

Thus, using (102) we have

$$\rho^{**} - \rho_0 = O_p(\lambda_n/s_n + \sqrt{n}\bar{x}_n^2/s_n^2 + \lambda_n/s_n) \xrightarrow{p} 0. \tag{106}$$

The latter convergence in (106) follows from the conditions (63). But (106) shows that the random variables ρ^{**} have the same limit as ρ_0 in (97).

So, the desired convergence (65) is proved.

Proof of Assertions from Example 4

First of all, note that

$$s_n^2 = \sum_{k=1}^{n} v_k^2 - \Big(\sum_{k=1}^{n} v_k\Big)/n \quad \text{for} \quad v_k = 1 - x_k = k^{-\alpha}. \tag{107}$$

Next, for $\alpha > 1/2$

$$\sum_{k=1}^{n} v_k^2 = \sum_{k=1}^{n} k^{-2\alpha} \sim n^{1-2\alpha}/(1-2\alpha),$$
$$\sum_{k=1}^{n} v_k = \sum_{k=1}^{n} k^{-\alpha} \sim n^{1-2\alpha}/(1-\alpha).$$

Thus, by (107)

$$s_n^2 \sim n^{1-2\alpha}\Big(\frac{1}{1-2\alpha} - \frac{1}{(1-\alpha)^2}\Big) = \frac{\alpha^2}{(1-2\alpha)(1-\alpha)^2} n^{1-2\alpha} \to \infty.$$

By this fact we obtain (71) since $\bar{x}_n \to 1$ and $\sum_{i=1}^{n} x_i^2 \sim n$ by (70).

Proofs of Assertions from Examples 5 and 6

It follows from (58) that

$$1 + \delta_1 = e^{\beta} e^{\tilde{\delta}_1} \quad \text{and} \quad \mathbf{E}\delta_1 = e^{\beta}\mathbf{E}^{\tilde{\delta}_1} - 1.$$

Thus, we have from (64) that $\mathbf{E}\delta_1 = 0$ and hence

$$e^{-\beta} = \mathbf{E}e^{\tilde{\delta}_1} \quad \text{and} \quad \text{var}\ \delta_1 = \mathbf{E}e^{2\tilde{\delta}_1}\Big/\Big(\mathbf{E}e^{\tilde{\delta}_1}\Big)^2 - 1. \tag{108}$$

So, if $\tilde{\delta}_1$ has a normal distribution with zero mean, then $\mathbf{E}e^{\lambda\tilde{\delta}_1} = e^{\lambda^2\tilde{\sigma}^2/2}$ and we have (72) since

$$\text{var}\ \delta_1 = e^{2^2\tilde{\sigma}^2/2}/(e^{\tilde{\sigma}^2/2})^2 - 1 = e^{\tilde{\sigma}^2} - 1 > \tilde{\sigma}^2.$$

In the case when (73) is true we have that

$$\mathbf{E}e^{\tilde{\delta}_1} = \cosh\tilde{\sigma}, \qquad \mathbf{E}e^{2\tilde{\delta}_1} = \cosh(2\tilde{\sigma}) = \cosh^2\tilde{\sigma} + \sinh^2\tilde{\sigma}.$$

These facts and (108) yield (74).

References

1. Sakhanenko, A.I., Linke, Y.Y.: Asymptotically normal estimation of a parameter in a linear-fractional regression problem. Sib. Math. J. **41**, 125–137 (2000)
2. Demidenko, E.: Optimization and Regression. Nauka, Moscow (1989)
3. Yermolenko, X.V., Sakhanenko, A.I.: Explicit asymptotically normal estimators of an unknown parameter in a partially linear regression problem. Sib. Electron. Math. Reports **10**, 719–726 (2013)
4. Savinkina, E.N., Sakhanenko, A.I.: Explicit estimators of an unknown parameter in a power regression problem. Sib. Electron. Math. Reports **11**, 725–733 (2014)
5. Kalenchuk, A.A., Sakhanenko, A.I.: The existence of explicit asymptotically normal estimators of an unknown parameter in a logarithmic regression problem. Sib. Electron. Math. Reports **12**, 874–883 (2015)
6. Heyde, C.C.: Quasy-Likelihood and Its Application: A General Approach to Optimal Parameter Estimation. Springer, New York (1997)
7. Seber, G.A.F., Wild, C.J.: Nonlinear Regression. Wiley, New York (2003)
8. Janssen, P., Jureckova, J., Veraverbeke, N.: Rate of convergence of one- and two-step M-estimators with applications to maximum likelihood and Pitman estimators. Ann. Statist. **13**, 1222–1229 (1985)
9. Verrill, S.: Rate of convergence of k-step Newton estimators to efficient likelihood estimators. Stat. Probab. Lett. **77**, 1371–1376 (2007)
10. Sakhanenko, A.I., Linke, Y.Y.: Asymptotically normal estimation of a multidimensional parameter in the linear-fractional regression problem. Sib. Math. J. **42**, 317–331 (2001)
11. Sakhanenko, A.I., Linke, Y.Y.: Asymptotically normal explicit estimation of parameters in the Michaelis – Menten equation. Sib. Math. J. **42**, 517–536 (2001)
12. Michaelis, L., Menten, M.L.: Die Kinetik der Invertinwirkung. Biochem Z **49**, 333–369 (1913)
13. Cornish-Bowden, A.: Fundamentals of Enzyme Kinetics. Butterworth, London (1979)
14. Ovcharenko, S.S., Sakhanenko, A.I.: Inconsistency of estimator of Johansen and Lumry for parameters of Michaelis - Menten equation. Vestnik Yugra State Univ. **8**, 95–99 (2008)

On the Behavior of the Risk of a LASSO-Type Estimator

Silvelyn Zwanzig and M. Rauf Ahmad

Abstract We introduce a LASSO-type estimator as a generalization of the classical LASSO estimator for non-orthogonal design. The generalization, named the SVD-LASSO, allows the design matrix to be of less than full rank. We assume fixed design matrix and normality but otherwise the properties of the SVD-LASSO do not necessarily rest on any strong conditions, particularly sparsity. We derive exact expressions for the risk of the SVD-LASSO and compare it with that of the corresponding ridge estimator.

Keywords Shrinkage estimation · High-dimensional inference · Linear models · SVD · MSE

1 Introduction

In the context of the theory of inference for the general linear model, the LASSO estimator is one of the most frequently used alternatives to the classical least-squares theory. Introduced by [1], the LASSO offers a regularized least-squares estimator, thus setting itself in competition with the ridge estimator [2, 3]. Deviating from each other on the basis of apparently a small difference, namely that the LASSO replaces the $\mathbb{L}_2$ penalty of ridge estimation with a $\mathbb{L}_1$ penalty on the unknown parameter vector to be estimated, the two shrinkage estimation methods lead to seriously different consequences. Probably the most attractive feature of LASSO is its simultaneous estimation-and-selection property lacked by its competitors. This, however, is also the most crucial aspect of LASSO that lends itself to critical evaluation.

Whereas ridge estimator is based on a simple algebraic solution to an ill-posed problem (hence equivalent to Tikhonov regularization in the general framework of

S. Zwanzig
Department of Mathematics, Uppsala University, Uppsala, Sweden
e-mail: zwanzig@math.uu.se

M.R. Ahmad (✉)
Department of Statistics, Uppsala University, Uppsala, Sweden
e-mail: rauf.ahmad@statistik.uu.se

J. Antoch et al. (eds.), *Analytical Methods in Statistics*, Springer Proceedings in Mathematics & Statistics 193, DOI 10.1007/978-3-319-51313-3_9

optimization theory), the LASSO estimator takes its strongest support from the bifurcation of the estimated parameter vector into an active set and the rest that is virtually forced to vanish by the edges of the $\mathbb{L}_1$ ball. The active set thus simultaneously does dimension reduction and makes LASSO applicable in high-dimensional settings. On the other hand, whereas the ridge estimator is computationally easily amenable using a few simple tools of linear algebra, the $\mathbb{L}_1$ penalty makes the LASSO problem mathematically intractable at several important fronts.

The point that we are interested to focus in this manuscript pertains to the aforementioned applicability of LASSO in high-dimensional or, more contextually speaking, sparse settings. As witnessed from the literature, the treatment of LASSO problem is almost invariably subjected to sparsity conditions under which the estimator is computed and its properties are studied. The sparsity condition comes in various ways with as many connotations, e.g., as sparsistency condition or restricted eigenvalue condition [4, 5].

Intrigued by this special feature of the LASSO, we are interested to evaluate its properties without particularly resorting much to the inevitability of the sparsity condition. In this context, we begin in the present study by introducing a LASSO-type estimator through an orthogonal transformation, assuming the design matrix to be fixed. A more detailed treatment of original, nontransformed, LASSO, for both fixed and random designs, is adjourned for further work.

2 Model, Assumptions and Some Basic Results

Consider the general linear model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{1}$$

with design matrix $\mathbf{X}_{n\times p}$, response vector $\mathbf{Y}_{n\times 1}$, parameter vector $\boldsymbol{\beta}_{p\times 1} = (\beta_1, \ldots, \beta_p)^T$ and the error vector $\boldsymbol{\varepsilon}_{n\times 1}$. The rows of $\mathbf{X}$, $\mathbf{X}_i = (X_{i1}, \ldots, X_{ip})'$, $i = 1, \ldots, n$, consist of p concomitant variables which can be categorial or continuous or mixed. We write $\mathbf{X} = (\mathbf{X}_{(1)}, \ldots, \mathbf{X}_{(p)})$ with $\mathbf{X}_{(j)} : n \times 1$ as jth column of $\mathbf{X}$, $j = 1, \ldots, p$. With $\mathbf{X}$ fixed, we set the following assumptions on $\boldsymbol{\varepsilon}$.

(A1) $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$, $\mathrm{Cov}(\boldsymbol{\varepsilon}) = \mathbf{I}$.

For inferential purposes, the model will be subjected to further assumptions in the sequel. Given the objectives detailed in Sect. 1 above, we also have $p \gg n$ with

$$r(\mathbf{X}) = k \leq n \ll p, \tag{2}$$

where $r(\cdot)$ denotes the rank of a matrix. The most preferred way of estimating $\boldsymbol{\beta}$ under A1, particularly without resorting to any distributional assumption, is through the well-known least-squares estimation (LSE). The LS estimator is defined as

$$\widehat{\beta}_{\text{LSE}} \in \arg\min_{\beta \in \mathbb{R}^p} \|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|^2 ,$$

where $\|\mathbf{a}\|^2 = \mathbf{a}'\mathbf{a}$ is the squared $\mathbb{L}_2$-norm of $\mathbf{a}$ (see [6], Chaps. 2–3 for details). This gives the solution set of the normal equations $\mathbf{X}^T\mathbf{Y} = \mathbf{X}^T\mathbf{X}\boldsymbol{\beta}$ as

$$\left\{ \left(\mathbf{X}^T\mathbf{X}\right)^- \mathbf{X}^T\mathbf{Y} + \mathbf{Q}\mathbf{z} \ : \ \mathbf{z} \in \mathbb{R}^p \right\},$$

where $\mathbf{Q} = \mathbf{I}_p - (\mathbf{X}^T\mathbf{X})^-(\mathbf{X}^T\mathbf{X})$ is the projection from $\mathbb{R}^p$ to the orthogonal space of $\mathscr{L}(\mathbf{X}^T\mathbf{X})$, and $(\cdot)^-$ denotes a g-inverse [7]. The minimum-variance estimator follows then at $\mathbf{z} = \mathbf{0}$ as

$$\widetilde{\beta} = (\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T\mathbf{Y}. \tag{3}$$

In the classical case where $p < n$ and $\mathbf{X}$ is of full rank so that $r(\mathbf{X}^T\mathbf{X}) = p$, $(\cdot)^-$ is replaced with the regular inverse $(\cdot)^{-1}$ leading to the unique solution $\widehat{\boldsymbol{\beta}}_{\text{LSE}}$. Note that, by the invariance of the projection $\mathbf{P} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T$ to g-inverse, the predicted vector $\mathbf{PY} = \widehat{\mathbf{Y}}$ is always unique even when $\widehat{\boldsymbol{\beta}}_{\text{LSE}}$ is not.

When $\mathbf{X}^T\mathbf{X}$ is ill-conditioned, e.g., under multicollinearity, with a special case of $\mathbf{X}$ being not of full rank, [3] introduced ridge estimator, $\widehat{\boldsymbol{\beta}}_{\text{Ridge}}$. Given $\lambda > 0$, it minimizes a penalized LS objective function,

$$\widehat{\boldsymbol{\beta}}_{\text{Ridge}} = \arg\min_{\boldsymbol{\beta} \in \mathbb{R}^p} \left(\|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda \|\boldsymbol{\beta}\|^2\right),$$

so that the normal equations $\mathbf{X}^T\mathbf{Y} = \left(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I}_p\right)\boldsymbol{\beta}$ are solved for

$$\widehat{\boldsymbol{\beta}}_{\text{Ridge}} = \left(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I}_p\right)^{-1} \mathbf{X}^T\mathbf{Y}.$$

Ridge regression uses $\mathbb{L}_2$-norm penalty, $\|\boldsymbol{\beta}\|^2$, and results into inflating the diagonal of the ill-conditioned $\mathbf{X}^T\mathbf{X}$ by a spherical matrix to guarantee its regular inverse. The ridge estimator reduces to $\widetilde{\boldsymbol{\beta}}$ for $\lambda = 0$. For some recent studies on ridge estimator, see e.g., [8, 9]. As another alternative to the LSE, [1] introduced LASSO which uses $\mathbb{L}_1$ norm of $\boldsymbol{\beta}$, i.e., $\|\boldsymbol{\beta}\|_1$ ([6], Chaps. 2–3), as penalty, so that

$$\widehat{\boldsymbol{\beta}}_{\text{LASSO}} \in \arg\min_{\boldsymbol{\beta} \in \mathbb{R}^p} \left(\|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda\|\boldsymbol{\beta}\|_1\right). \tag{4}$$

The LASSO estimator is a solution to the Karush–Kuhn–Tucker (KKT) conditions

$$\mathbf{X}^T\mathbf{Y} - \mathbf{X}^T\mathbf{X}\boldsymbol{\beta} = \lambda\omega \quad \text{or} \quad \mathbf{X}_{(j)}^T\mathbf{Y} - \mathbf{X}_{(j)}^T\mathbf{X}\boldsymbol{\beta}_j = \lambda\omega_j, \ \ j = 1, \ldots, p,$$

with ω_j as defined below where $\text{sgn}(\cdot)$ denotes the sign function

$$\omega_j \in \begin{cases} \text{sgn}(\beta_j) \text{ if } \beta_j \neq 0 \\ [-1, 1] \ \text{ if } \beta_j = 0. \end{cases}$$

A unique feature of LASSO, and one of the foremost reasons of its use in practice, is that it does model estimation and selection simultaneously, hence the name (least absolute shrinkage and selection operator). This is a consequence of the $\mathbb{L}_1$ penalty which forces several $\widehat{\beta}_j$'s to be exactly zero. Thus, the target of LASSO is the set of nonzero estimators, the so-called active set, defined as

$$S = \left\{ j : \left| \mathbf{X}_{(j)}^T (\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}_{\text{LASSO}}) \right| = \lambda \right\}.$$

Now, let $\mathbf{X}_S = (\mathbf{X}_{(j)})_{j \in S}$ denote the $n \times s$ design matrix corresponding to the set S with its cardinality $s = \#(S)$ and the $s \times 1$ vector sgn $=$ sgn $\left(\mathbf{X}_S^T (\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}_{\text{LASSO}})\right)$, where $\widehat{\beta}_{\text{LASSO},j} = 0$ for $j \notin S$. When $r(\mathbf{X}_S) = s$, we can solve the KKT

$$\mathbf{X}_S^T (\mathbf{Y} - \mathbf{X}_S \widehat{\boldsymbol{\beta}}_{\text{LASSO},S}) = \lambda \text{ sgn} \;\Rightarrow\; \widehat{\boldsymbol{\beta}}_{\text{LASSO},S} = \left(\mathbf{X}_S^T \mathbf{X}_S\right)^{-1} (\mathbf{X}_S^T \mathbf{Y} - \lambda \text{ sgn}),$$

where $\widehat{\boldsymbol{\beta}}_{\text{LASSO},S}$ denotes the nonzero estimators corresponding to the set S, so that $\widehat{\boldsymbol{\beta}}_{\text{LASSO},-S} = 0$, where $-S$ denotes the complementary set to S. Note that, the LASSO estimator is not invariant, not even under an orthogonal transformation. In the next section, we introduce a LASSO-type estimator using an orthogonal transformation, i.e., a two-step estimator where the first step decomposes the design matrix through its SVD and the second step consists of LASSO-estimation of the transformed model. We study the properties of this estimator, particularly its risk, and compare it with the corresponding ridge estimator.

3 The SVD-LASSO

Consider Model (1) with $\mathbf{X}$ fixed. Given a singular value decomposition of $\mathbf{X}$

$$\mathbf{X} = \mathbf{U}\mathbf{L}\mathbf{V}^T \tag{5}$$

where $\mathbf{U} : n \times n$ and $\mathbf{V} : p \times p$ are composed of orthonormal eigenvectors of $\mathbf{X}\mathbf{X}^T$ and $\mathbf{X}^T\mathbf{X}$, respectively, so that $\mathbf{U}^T\mathbf{U} = \mathbf{I}_n$, $\mathbf{V}^T\mathbf{V} = \mathbf{I}_p$, and $\mathbf{L} = \text{diag}(\lambda_1, \ldots, \lambda_k, 0, \ldots, 0)$ is $n \times p$ diagonal matrix with $\lambda_1^2, \ldots, \lambda_k^2$ as nonzero eigenvalues of $\mathbf{X}\mathbf{X}^T$ (or of $\mathbf{X}^T\mathbf{X}$) or, correspondingly, $\lambda_1, \ldots, \lambda_k$ as the singular values. Define

$$\mathbf{Z} = \mathbf{U}\mathbf{L}, \tag{6}$$

so that

$$\mathbf{Z}^T\mathbf{Z} = \mathbf{L}^T\mathbf{L} = \text{diag}(\lambda_1^2, \ldots, \lambda_k^2, 0, \ldots, 0) \tag{7}$$

is the $p \times p$ diagonal matrix. Moreover, writing $\mathbf{X}\boldsymbol{\beta} = \mathbf{Z}\boldsymbol{\alpha}$, we have

$$\mathbf{U}\mathbf{L}\mathbf{V}^T\boldsymbol{\beta} = \mathbf{U}\mathbf{L}\boldsymbol{\alpha} \;\Rightarrow\; \boldsymbol{\alpha} = \mathbf{V}^T\boldsymbol{\beta} \;\Rightarrow\; \boldsymbol{\beta} = \mathbf{V}\boldsymbol{\alpha}. \tag{8}$$

The transformed model can now be stated as

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \tag{9}$$

with the LASSO estimator of $\boldsymbol{\alpha}$ defined as

$$\widehat{\boldsymbol{\alpha}} \in \arg\min F(\boldsymbol{\alpha}) = \arg\min \left(\|\mathbf{Y} - \mathbf{Z}\boldsymbol{\alpha}\|^2 + \lambda\|\boldsymbol{\alpha}\|_1\right). \tag{10}$$

The corresponding estimator under the transformation, heretofore called the SVD-LASSO estimator, is obtained as $\widehat{\boldsymbol{\beta}} = \mathbf{V}\widehat{\boldsymbol{\alpha}}$. Following definition sums up both cases.

Definition 1 Consider Model (9) and the SVD in (5), where $\boldsymbol{\beta} = \mathbf{V}\boldsymbol{\alpha}$. Then, $\widehat{\boldsymbol{\alpha}}$ in (10) is the LASSO estimator and $\widehat{\boldsymbol{\beta}} = \mathbf{V}\widehat{\boldsymbol{\alpha}}$ is the corresponding SVD-LASSO estimator.

Recall that, for LS and Ridge estimators it holds, for any arbitrary orthogonal matrix $\mathbf{Q}$, that $\widehat{\boldsymbol{\beta}}_{\text{LSE}} = \mathbf{Q}\widehat{\boldsymbol{\alpha}}_{\text{LSE}}$ and $\widehat{\boldsymbol{\beta}}_{\text{Ridge}} = \mathbf{Q}\widehat{\boldsymbol{\alpha}}_{\text{Ridge}}$, where

$$\widehat{\alpha}_{\text{LSE}} \in \arg\min_{\boldsymbol{\alpha}\in\mathbb{R}^p} \|\mathbf{Y} - \mathbf{Z}\boldsymbol{\alpha}\|^2 \tag{11}$$

$$\widehat{\boldsymbol{\alpha}}_{\text{Ridge}} = \arg\min_{\boldsymbol{\alpha}\in\mathbb{R}^p} \left(\|\mathbf{Y} - \mathbf{Z}\boldsymbol{\alpha}\|^2 + \lambda\, \|\boldsymbol{\alpha}\|^2\right). \tag{12}$$

But the same does not hold for LASSO. By Definition 1, it, however, does hold for the proposed SVD-LASSO under the specific transformation given in (5), i.e., $\widehat{\boldsymbol{\beta}}_{\text{SVD-LASSSO}} = \mathbf{V}\widehat{\boldsymbol{\alpha}}_{\text{LASSO}}$; see also Theorem 2. Note also that, $\widehat{\boldsymbol{\alpha}} = \widehat{\boldsymbol{\beta}}$ if $\mathbf{V} = \mathbf{I}$.

We begin with our main results by stating the following theorem on some general identities that will be subsequently specialized. All proofs are in Appendix.

Theorem 1 *Given Model (1). Let* $\mathbf{W}$ *be an* $n \times n$ *positive semi-definite matrix and* $\mathbf{A}$ *be an arbitrary* $p \times p$ *matrix. Then, we have the following.*

(a) The objective function

$$F(\boldsymbol{\beta}) = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T\mathbf{W}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) + \lambda\|\mathbf{A}\boldsymbol{\beta}\|_1 \tag{13}$$

is convex and its sub-differential is the set

$$\partial F(\boldsymbol{\beta}) = \{-2\mathbf{X}^T\mathbf{W}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) + \lambda\mathbf{A}^T\boldsymbol{\omega} : \boldsymbol{\omega} \in \mathscr{L}(\mathbf{A}\boldsymbol{\beta})\}, \tag{14}$$

where $\mathscr{L}(\mathbf{b}) = \{\boldsymbol{\omega} : \omega_j = sgn(b_j) \text{ for } b_j \neq 0, \omega_j \in [-1, 1] \text{ for } b_j = 0\}$ *with* $\{\boldsymbol{\omega} = (\omega_1, \ldots, \omega_p)^T\}$*; we can write*

$$\omega_j(\mathbf{b}) = \begin{cases} -1 & \text{if } b_j < 0 \\ 1 & \text{if } b_j > 0 \\ a_j & \text{if } b_j = 0,\ a_j \in [-1, 1]. \end{cases} \tag{15}$$

(b) For $F(\boldsymbol{\beta})$ *in (13),* $\widehat{\boldsymbol{\beta}} \in \arg\min F(\boldsymbol{\beta}) \Leftrightarrow \exists\, a_j \in [-1, 1]$ *such that*

$$\mathbf{X}_{(j)}^T \mathbf{W} \mathbf{X}_{(j)} \widehat{\boldsymbol{\beta}} = \mathbf{X}_{(j)}^T \mathbf{W} \mathbf{Y} - \frac{\lambda}{2} \mathbf{A}^T \widehat{\boldsymbol{\omega}}_j, \tag{16}$$

with $\widehat{\boldsymbol{\omega}}_j = \omega_j(\mathbf{A}\widehat{\boldsymbol{\beta}})$ *as in (15).*

The following theorem formally juxtaposes the estimators given in Definition 1.

Theorem 2 *Let Models (1) and (9) hold with* $\boldsymbol{\alpha} = \mathbf{V}^T \boldsymbol{\beta}$. *Then* $\widehat{\boldsymbol{\alpha}} = \mathbf{V}^T \widehat{\boldsymbol{\beta}}$ *with*

$$\widehat{\boldsymbol{\alpha}} = \arg\min \left\{ \|\mathbf{Y} - \mathbf{Z}\boldsymbol{\alpha}\|^2 + \lambda \|\boldsymbol{\alpha}\|_1 \right\} \tag{17}$$

$$\widehat{\boldsymbol{\beta}} = \arg\min \left\{ \|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda \|\mathbf{V}^T \boldsymbol{\beta}\|_1 \right\}. \tag{18}$$

Benefiting from the results of Theorem 1, we now develop basic identities for the transformed model. For a more comprehensive treatment, we reduce all computations at the column level of the design matrix, i.e., $\mathbf{Z}_{(j)}$, $j = 1, \ldots, p$. Consider then

$$\mathbf{Z}_{(j)}^T \mathbf{Z} \widehat{\boldsymbol{\alpha}} = \mathbf{Z}_{(j)}^T \mathbf{Y} - \frac{\lambda}{2} \omega_j(\widehat{\boldsymbol{\alpha}}), \quad j = 1, \ldots, p, \tag{19}$$

where, for $j = 1, \ldots, k$,

$$\mathbf{Z}_{(j)}^T \mathbf{X} \boldsymbol{\beta} = \mathbf{Z}_{(j)}^T \mathbf{Z} \boldsymbol{\alpha} = \alpha_j \lambda_j^2 \text{ and } \mathbf{Z}_{(j)}^T \mathbf{Z}_{(j)} = \lambda_j^2, \tag{20}$$

so that, from (19), $\lambda_j^2 \widehat{\alpha}_j = \mathbf{Z}_{(j)}^T \mathbf{Y} - (\lambda/2)\omega_j(\widehat{\boldsymbol{\alpha}})$. Thus, for $j = 1, \ldots, k$,

$$\widehat{\alpha}_j > 0 \Leftrightarrow \lambda_j^2 \widehat{\alpha}_j > 0 \Rightarrow \mathbf{Z}_{(j)}^T \mathbf{Y} - \lambda/2 > 0 \Rightarrow \mathbf{Z}_{(j)}^T \mathbf{Y} > \lambda/2.$$

Likewise, $\widehat{\alpha}_j < 0 \Leftrightarrow \mathbf{Z}_{(j)}^T \mathbf{Y} < -\lambda/2$ and $\widehat{\alpha}_j = 0 \Leftrightarrow \|\mathbf{Z}_{(j)}^T \mathbf{Y}\| \leq \lambda/2$. Moreover, from (19) and $\mathbf{Z} = \mathbf{UL}$ in (6), it follows for $j = k+1, \ldots, p$ that $\mathbf{Z}_{(j)} = \mathbf{0} \Rightarrow 0 = 0 - (\lambda/2)\omega_j(\widehat{\boldsymbol{\alpha}})$ so that $\omega_j(\widehat{\boldsymbol{\alpha}}) = 0$ which implies $\widehat{\alpha}_j = 0$. Now, define

$$z_{(j)} = \lambda_j^{-2} \mathbf{Z}_{(j)}^T \mathbf{Y}, \quad j = 1, \ldots, k, \tag{21}$$

with $\mathrm{E}(z_{(j)}) = \alpha_j$, $\mathrm{Var}(z_{(j)}) = \lambda_j^{-2}$. Then we can write

$$\widehat{\alpha}_j = z_{(j)} - a_j(\lambda) \widehat{\omega}_j(\widehat{\boldsymbol{\alpha}}), \tag{22}$$

where

$$a_j(\lambda) = \lambda / 2\lambda_j^2. \tag{23}$$

Using these computations, we define the following conditional moments.

$$\mu_j(\lambda) = \mathrm{E}\left(z_{(j)} \mid \left|z_{(j)}\right| > a_j(\lambda)\right) \tag{24}$$

$$\sigma_j^2(\lambda) = \mathrm{Var}\left(z_{(j)} \mid \left|z_{(j)}\right| > a_j(\lambda)\right) \tag{25}$$

$$e_j(\lambda) = \mathrm{E}\left(z_{(j)} \mid z_{(j)} > a_j(\lambda)\right) - \mathrm{E}\left(z_{(j)} \mid z_{(j)} < -a_j(\lambda)\right). \tag{26}$$

Further, in this context, we define the following probabilities

$$p_{+j} = \mathrm{P}\left(z_{(j)} > a_{(j)}\right), \quad p_{-j} = \mathrm{P}\left(z_{(j)} < -a_{(j)}\right), \quad p_{0j} = \mathrm{P}\left(\left|z_{(j)}\right| < a_{(j)}\right), \quad (27)$$

which, along with (21)–(26) will help us prove several main results without depending on any serious distributional assumptions. In fact, this point needs to be emphasized more clearly before we proceed further. What this essentially means is that, (21)–(27) will suffice for us to prove all but only one or two of the results in the sequel without requiring normality of $z_{(j)}$ or any other similar distributional assumption. For this, we begin with the following result which summarizes $\widehat{\alpha}_j$ as a special consequence of Theorem 1.

Corollary 1 *Given Theorem 1(b). Let* $\mathbf{W} = \mathbf{I}_n$ *and* $\mathbf{A} = \mathbf{I}_p$. *Then (see 22)*

$$\widehat{\alpha}_j = z_{(j)} - a_j(\lambda)\widehat{\omega}_j, \quad j = 1, \ldots, k \quad (28)$$

$$= 0, \qquad j = k+1, \ldots, p. \quad (29)$$

Corollary 1 gives component-wise regularized estimators for Model (9). Considering the rank condition in (2) and noting that $r(\mathbf{X}) = r(\mathbf{Z}\mathbf{V}^T) = r(\mathbf{Z}) = r(\mathbf{Z}^T\mathbf{Z})$, Corollary 1 implies a flexibility in the form and rank of $\mathbf{Z}^T\mathbf{Z}$; e.g., we do not assume $\mathbf{Z}^T\mathbf{Z} = \mathbf{I}$ or any other structure that might limit the rank, hence the cardinality of the active set of $\boldsymbol{\alpha}_j$'s.

The following theorem uses the conditions on $\widehat{\alpha}_j$ to establish conditions on the bilinear form $z_{(j)} = \mathbf{Z}_{(j)}^T\mathbf{Y}$ in terms of tuning parameter λ. These conditions will drastically help us prove several results and study the risk of $\widehat{\boldsymbol{\alpha}}$ in Sect. 4. The proof of Theorem 3 follows from the computations given above and is therefore omitted.

Theorem 3 *Given Theorem 2 with* $\widehat{\boldsymbol{\alpha}} = \arg\min F(\boldsymbol{\alpha})$. *Then, for* $j = 1, \ldots, k$,

$$\widehat{\alpha}_j \begin{cases} > 0 \Leftrightarrow z_{(j)} > a_j \\ < 0 \Leftrightarrow z_{(j)} < -a_j \\ = 0 \Leftrightarrow |z_{(j)}| < a_j \end{cases} \quad (30)$$

with $z_{(j)}$ *in (21) and* a_j *in (23). Further,* $\widehat{\alpha}_j = 0$ *for* $j = k+1, \ldots, p$.

4 Computation of the Risk

In this section, we study mean-squared error of $\widehat{\boldsymbol{\alpha}}$ and compare it with that of ridge estimator. First note that $\|\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\|^2 = \|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\|^2$ by the orthogonality of $\mathbf{V}$ since MSE is invariant under the orthogonal transformation, i.e., MSE($\widehat{\boldsymbol{\alpha}}$) = MSE($\widehat{\boldsymbol{\beta}}$). That is, MSEs of estimated parameter vectors for Models (1) and (9) are same under the transformation, even if the estimators themselves are not. Now, recall that,

$$\mathrm{MSE}(\widehat{\theta}) = \mathrm{E}\|\widehat{\theta} - \theta\|^2 = [\mathrm{bias}(\widehat{\theta})]^2 + \mathrm{Var}(\widehat{\theta}) = \|\mathrm{E}(\widehat{\theta}) - \theta\|^2 + \mathrm{E}\|\widehat{\theta} - \mathrm{E}(\widehat{\theta})\|^2,$$

for any estimator $\widehat{\theta}$ of θ, where $\mathrm{E}(\widehat{\theta}) - \theta = \mathrm{bias}(\widehat{\theta})$. In the following, Theorem 4 gives the moments of $\widehat{\alpha}_j$ under general conditions and Corollary 2 specializes these results to the normal case and adds the corresponding expressions for MSE.

Theorem 4 *Given (23)–(27) and* $\widehat{\boldsymbol{\alpha}}$ *as in Theorem 2 with* $\widehat{\boldsymbol{\alpha}} = \mathbf{V}^T\widehat{\boldsymbol{\beta}}$. *Then*

$$E\left(\widehat{\alpha}_j\right) = \mu_j(\lambda) - a_j(\lambda)\frac{p_{+j} - p_{-j}}{1 - p_{0j}} \quad j = 1, \ldots, k, \tag{31}$$

$$Var(\widehat{\alpha}_j) = \sigma_j^2(\lambda) + 2a_j(\lambda)\frac{p_{+j}p_{-j}}{(1 - p_{0j})^2}\left\{2a_j(\lambda) - e_j(\lambda)\right\} \quad j = 1, \ldots, k. \tag{32}$$

$$[bias(\widehat{\boldsymbol{\alpha}})]^2 = \sum_{j=1}^{k}\left[bias(\widehat{\alpha}_j)\right]^2 + \sum_{j=k+1}^{p}\alpha_j^2 \tag{33}$$

$$MSE(\widehat{\boldsymbol{\beta}}) = MSE(\widehat{\boldsymbol{\alpha}}) = \sum_{j=1}^{k} E\left(\widehat{\alpha}_j - \alpha_j\right)^2 + \sum_{j=k+1}^{p}\alpha_j^2, \tag{34}$$

with $E\left(\widehat{\alpha}_j - \alpha_j\right)^2 = Var(\widehat{\alpha}_j) + \left[bias(\widehat{\alpha}_j)\right]^2$ *and* $bias(\widehat{\alpha}_j) = E\left(\widehat{\alpha}_j\right) - \alpha_j$.

Corollary 2 *Given Theorem 4 and let the normality assumption holds. Then*

$$E\left(\widehat{\alpha}_j\right) = \alpha_j(1 - p_{0j}) + \lambda_j^{-2}d_j - a_j(\lambda)\left(p_{+j} - p_{-j}\right) \quad j = 1, \ldots, k, \tag{35}$$

$$Var(\widehat{\alpha}_j) = \frac{1}{\lambda_j^2}\left[1 - p_{0j} + v_j + \frac{1}{1 - p_{0j}}d_j^2 + 2\lambda\frac{p_{+j}p_{-j}}{1 - p_{0j}}\left(a_j(\lambda) - e_j\right)\right] \tag{36}$$

$$[bias(\widehat{\boldsymbol{\alpha}})]^2 = \sum_{j=1}^{k}\left(\alpha_j p_{0j} - \lambda_j^{-2}d_j + a_j(\lambda)(p_{+j} - p_{-j})\right) + \sum_{j=k+1}^{p}\alpha_j^2 \tag{37}$$

$$MSE(\widehat{\boldsymbol{\alpha}}) = \sum_{j=1}^{k}\left(Var(\widehat{\alpha}_j) + \left[bias(\widehat{\alpha}_j)\right]^2\right) + \sum_{j=k+1}^{p}\alpha_j^2 \tag{38}$$

with $bias(\widehat{\alpha}_j) = \alpha_j p_{0j} + a_j(\lambda)(p_{+j} - p_{-j}) - \lambda_j^{-2}d_j$. *Further,* $d_j = d_j(\lambda, \lambda_j, \alpha)$, $v_j = d_j(\lambda, \lambda_j, \alpha)$ *and* $e_j = d_j(\lambda, \lambda_j, \alpha)$ *are defined as following.*

$$d_j = \varphi_{(\alpha_j, \lambda_j^{-2})}\left(a_j(\lambda)\right) - \varphi_{(\alpha_j, \lambda_j^{-2})}\left(-a_j(\lambda)\right) \tag{39}$$

$$v_j = \left(a_j(\lambda) - \alpha_j\right)\varphi_{(\alpha_j, \lambda_j^{-2})}\left(a_j(\lambda)\right) + \left(a_j(\lambda) + \alpha_j\right)\varphi_{(\alpha_j, \lambda_j^{-2})}\left(-a_j(\lambda)\right) \tag{40}$$

$$e_j = \frac{1}{\lambda_j^2}\left(\frac{1}{p_{+,j}}\varphi_{(\alpha_j, \lambda_j^{-2})}\left(a_j(\lambda)\right) - \frac{1}{p_{-,j}}\varphi_{(\alpha_j, \lambda_j^{-2})}\left(-a_j(\lambda)\right)\right). \tag{41}$$

Before we evaluate the risk of LASSO estimator, we develop the same for ridge estimator in order to compare the two later. For this, consider the model

$$\begin{pmatrix}\mathbf{Y}\\ \mathbf{0}\end{pmatrix} = \begin{pmatrix}\mathbf{X}\\ \sqrt{\lambda}\mathbf{B}\end{pmatrix}\boldsymbol{\beta} + \begin{pmatrix}\boldsymbol{\varepsilon}\\ \boldsymbol{\delta}\end{pmatrix}$$

with $\mathbf{B}$ and $\boldsymbol{\delta}$ of appropriate order so that the LSE criterion gives

$$\left\| \begin{pmatrix} \mathbf{Y} \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} \mathbf{X} \\ \sqrt{\lambda}\mathbf{B} \end{pmatrix} \boldsymbol{\beta} \right\|^2 = \|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda \|\mathbf{0} - \mathbf{B}\boldsymbol{\beta}\|^2$$

and we can write

$$\widehat{\boldsymbol{\beta}}_{\text{Ridge}} = \left(\begin{pmatrix} \mathbf{X} \\ \sqrt{\lambda}\mathbf{B} \end{pmatrix}^T \begin{pmatrix} \mathbf{X} \\ \sqrt{\lambda}\mathbf{B} \end{pmatrix} \right)^- \begin{pmatrix} \mathbf{X} \\ \sqrt{\lambda}\mathbf{B} \end{pmatrix}^T \begin{pmatrix} \mathbf{Y} \\ \mathbf{0} \end{pmatrix} = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{B}^T\mathbf{B})^- \mathbf{X}^T\mathbf{Y} \quad (42)$$

and $\mathbf{X}\widehat{\boldsymbol{\beta}}_{\text{Ridge}} = \mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{B}^T\mathbf{B})^-\mathbf{X}^T\mathbf{Y}$, where $\widehat{\boldsymbol{\beta}}_{\text{Ridge}}$ and $\mathbf{X}\widehat{\boldsymbol{\beta}}_{\text{Ridge}}$ reduce to the usual ridge estimation and prediction when $\mathbf{B} = \mathbf{I}$, the case that will be followed in the sequel. Further, in this case

$$\mathbf{X}\widehat{\boldsymbol{\beta}}_{\text{Ridge}} = \mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I}_p)^{-1}\mathbf{X}^T\mathbf{Y} = \mathbf{Z}(\mathbf{Z}^T\mathbf{Z} + \lambda\mathbf{V}\mathbf{V}^T)^{-1}\mathbf{Z}^T\mathbf{Y} = \mathbf{Z}\widehat{\boldsymbol{\alpha}}_{\text{Ridge}}.$$

Note that, the MSE of ridge estimator is invariant under any arbitrary orthogonal transformation, i.e., $\text{MSE}(\widehat{\boldsymbol{\alpha}}_{\text{Ridge}}) = \text{MSE}(\widehat{\boldsymbol{\beta}}_{\text{Ridge}})$. The following theorem summarizes the results for ridge estimator.

Theorem 5 *Let $\widehat{\boldsymbol{\alpha}}_{Ridge} = \mathbf{V}^T\widehat{\boldsymbol{\beta}}_{Ridge}$ be the ridge estimator where $\widehat{\boldsymbol{\beta}}_{Ridge}$ is as defined in (42) with $\mathbf{B} = \mathbf{I}$. Then*

$$E\left(\widehat{\boldsymbol{\alpha}}_{Ridge,j}\right) = \frac{\lambda_j^2}{\lambda_j^2 + \lambda}\alpha_j \quad j = 1, \ldots, k, \quad (43)$$
$$= 0 \qquad j = k+1, \ldots, p.$$

$$\left[bias\left(\widehat{\boldsymbol{\alpha}}_{Ridge}\right)\right]^2 = \sum_{j=1}^{k}\left(\frac{\alpha_j\lambda}{\lambda_j^2 + \lambda}\right)^2 + \sum_{j=k+1}^{p}\alpha_j^2 \quad (44)$$

$$Var\left(\widehat{\boldsymbol{\alpha}}_{Ridge}\right) = \sum_{j=1}^{k}\left(\frac{\lambda_j}{\lambda_j^2 + \lambda}\right)^2 \quad (45)$$

$$MSE\left(\widehat{\boldsymbol{\alpha}}_{Ridge}\right) = \sum_{j=1}^{k}\left[\left(\frac{\alpha_j\lambda}{\lambda_j^2 + \lambda}\right)^2 + \left(\frac{\lambda_j}{\lambda_j^2 + \lambda}\right)^2\right] + \sum_{j=k+1}^{p}\alpha_j^2. \quad (46)$$

Further, the limiting values of MSE of $\widehat{\boldsymbol{\alpha}}_{Ridge}$ are given as

$$\lim_{\lambda \to 0} MSE(\widehat{\boldsymbol{\alpha}}_{Ridge}) = \sum_{j=1}^{k}\frac{1}{\lambda_j^2} + \sum_{j=k+1}^{p}\alpha_j^2 \quad (47)$$

$$\lim_{\lambda \to \infty} MSE(\widehat{\boldsymbol{\alpha}}_{Ridge}) = \sum_{j=1}^{p}\alpha_j^2. \quad (48)$$

From the preceding computations, it is obvious that the bias and variance of LASSO estimator $\widehat{\boldsymbol{\alpha}}$ are functions of both the true parameter, $\boldsymbol{\alpha}$, and the regularization parameter, λ, so that the behavior of the risk must be studied in terms of both arguments. The following theorem summarizes important properties of the risk of $\widehat{\boldsymbol{\alpha}}$. Note that, like Corollary 2, Theorem 6 also requires normality assumption.

Theorem 6 *Let* $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \mathbf{I})$. *The following holds for bias and MSE of* $\widehat{\boldsymbol{\alpha}}$.

I. For every $\lambda > 0$, $bias(\widehat{\boldsymbol{\alpha}}_j) = E_\alpha(\widehat{\alpha}_j) - \alpha_j$ *is an odd function with respect to* α_j, *i.e.,* $bias(\alpha_j) = -bias(-\widehat{\alpha}_j)$, $j = 1, \ldots, k$.

II. For every $\lambda > 0$, $\left[bias(\widehat{\boldsymbol{\alpha}}_j)\right]^2$ *is a monotonically increasing function with respect to* α_j, $j = 1, \ldots, k$.

III. For every $\alpha > 0$, $[bias(\widehat{\boldsymbol{\alpha}})]^2 = \sum_{j=1}^k [bias(\widehat{\alpha}_j)]^2$ *is a monotonically increasing in* λ.

IV. For the variance as function of λ *with true parameter vanishing, i.e.,* $Var_0(\lambda) := \sum_{j=1}^k Var(\widehat{\alpha}_j)$ *for* $\alpha_j = 0$, $j = 1, \ldots, k$, *we have the following:* $\exists\, \lambda_0 > 0$ *such that*

$$\frac{d}{d\lambda} Var_0(\lambda) = \begin{cases} > 0 & for\ \lambda < \lambda_0 \\ = 0 & for\ \lambda = \lambda_0 \\ < 0 & for\ \lambda > \lambda_0. \end{cases} \tag{49}$$

V. For every α, *we have the following limits of MSE*($\widehat{\boldsymbol{\alpha}}$).

$$\lim_{\lambda \to 0} MSE(\widehat{\boldsymbol{\alpha}}) = \sum_{j=1}^{k} \frac{1}{\lambda_j^2} + \sum_{j=k+1}^{p} \alpha_j^2 \tag{50}$$

$$\lim_{\lambda \to \infty} MSE(\widehat{\boldsymbol{\alpha}}) = \sum_{j=1}^{p} \alpha_j^2. \tag{51}$$

Discussion, Comparisons, and Conclusions

This paper presents a LASSO-type estimator using an orthogonal transformation through the SVD of the design matrix. To study its risk behavior, the mean-squared error of the proposed estimator is computed, its properties are theoretically evaluated, and compared with the risk of corresponding ridge estimator.

Whereas the MSE of LASSO and ridge estimators approach the same values in the limit (Theorem 6), the differential features of the two estimators need a comment

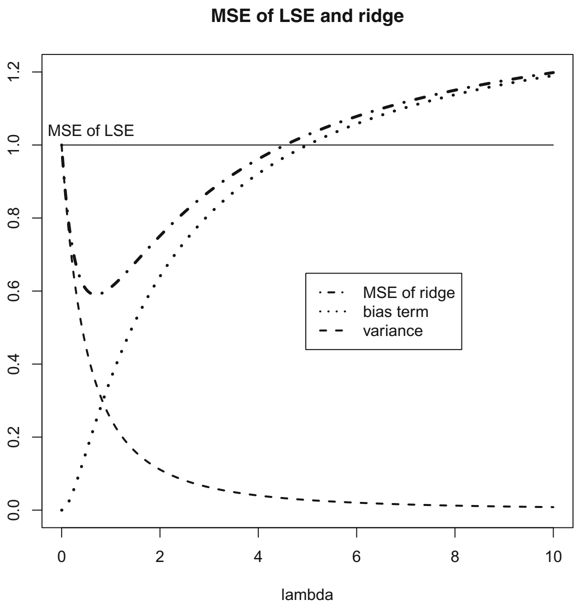


Fig. 1 Effect of tuning parameter λ on the risk of Ridge estimator

or two. First, as already mentioned around Definition 1, the ridge estimator is invariant under any arbitrary orthogonal transformation, whereas the proposed LASSO estimator depends on the special orthogonal transformation under the SVD of the design matrix as given in (5).

Further, concerning the risk of two estimators as function of the tuning parameter, we first notice that the variance of the ridge estimator does not depend on the true parameter. We then observe that the bias of ridge estimator increases with λ whereas the variance decreases. We also notice an improvement in the estimator for small λ values. The bias for LASSO estimator also increases, but the variance has opposite trend for small true parameter values (first increasing then decreasing), where it improves for large λ values. Moreover, its variance increases when considered with a combination of small λ values and small values of the true parameter $\boldsymbol{\alpha}$ (See also Figs. 1 and 2).

Asymptotic MSE for two extreme cases is also studied. When $\lambda \to 0$, the variance reduces to be inversely proportional to the nonzero eigenvalues, λ_j, of $\mathbf{X}'\mathbf{X}$ where bias becomes directly proportional to the squared true parameters, so that it vanishes when the true parameter is zero. On the other hand, for $\lambda \to \infty$, the variance vanishes completely and the entire MSE is formulated by the bias component, now $\|\boldsymbol{\alpha}\|^2$. Although we have not dedicated much space to the case of LSE estimator, particularly its comparison to the proposed LASSO estimator, but it can be verified that the results reduce to those of LSE, $\widetilde{\boldsymbol{\beta}}$, under respective conditions. For example, the MSE of LASSO, Ridge and LSE $\widetilde{\boldsymbol{\beta}}$ are all same for $\lambda = 0$.

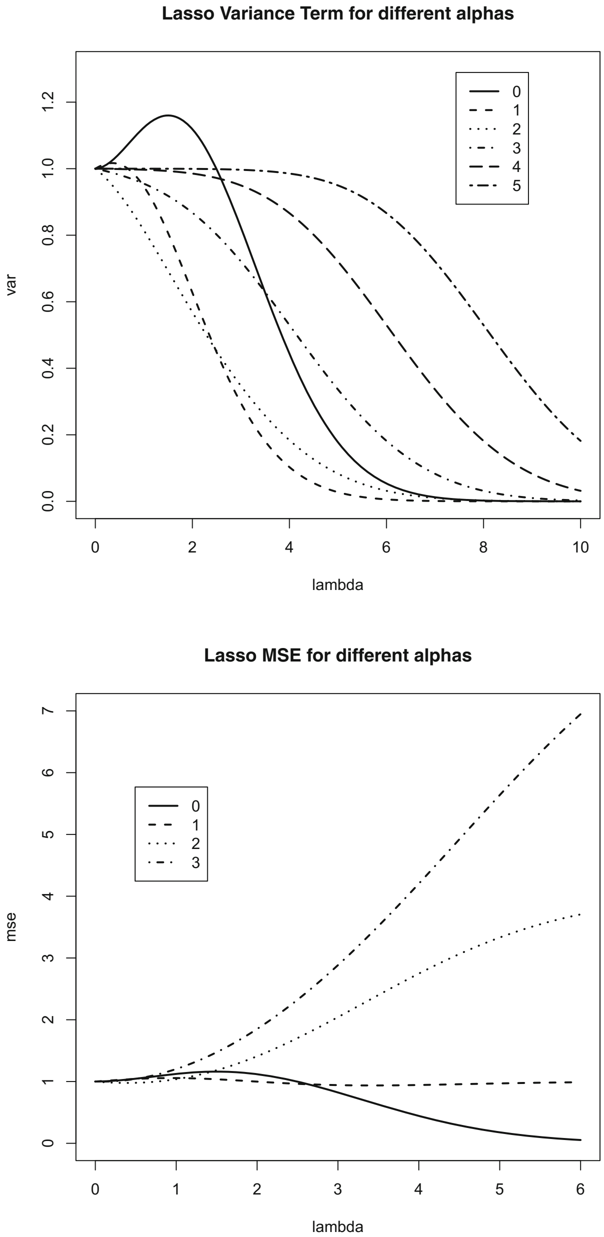


Fig. 2 Effect of tuning parameter λ on the risk of LASSO estimator

Acknowledgements The authors are thankful to the reviewer for several constructive comments which lead to an improved version of the article.

Appendix

4.1 Two Basic Results

We begin by stating two fundamental results that are essentially needed to prove the main theorems. The first of these (Lemma 1) is on the covariance between a continuous random variable and a discrete random variable with two-point distribution. The second result (Theorem 7) provides mean and variance of a truncated distribution. For basic introduction to truncated distributions, see [10].

Lemma 1 *Let X be a continuous random variables with $E(X) = \mu$. Let Y be a discrete random variable following a two-point distribution with support $\{-1, 1\}$, where $P(Y = 1) = p$. Then*

$$Cov(X, Y) = 2p(1-p)[E(X \mid Y = 1) - E(X \mid Y = -1)]. \tag{52}$$

Proof First note that $\mathrm{E}(Y) = 2p - 1$. Now, by definition,

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= \mathrm{E}(X - \mu)(Y - 2p + 1) = \mathrm{E}[Y(X - \mu)] \\
&= \mathrm{E}[Y(X - \mu) \mid Y = -1](1 - p) + \mathrm{E}[(Y(X - \mu) \mid Y = 1)]p \\
&= (1 - 2p)\mu - \mathrm{E}(X \mid Y = -1)(1 - p) + \mathrm{E}(X \mid Y = 1)p.
\end{aligned}$$

Since $\mu = EX = E(X \mid Y = -1)(1 - p) + E(X \mid Y = 1)p$, it follows that

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= (1 - 2p)\mathrm{E}(X \mid Y = -1)(1 - p) + (1 - 2p)\mathrm{E}(X \mid Y = 1)p \\
&= 2p(1 - p)[\mathrm{E}(X \mid Y = 1) - \mathrm{E}(X \mid Y = -1)].
\end{aligned}$$

□

Theorem 7 *Let $X \sim N(\mu, \sigma^2)$ with its pdf $\varphi_{(\mu,\sigma^2)}(x)$. Let $A = \{X \mid |X| > a\}$, $a \in \mathbb{R}$. The moments of the truncated distribution of X given A are*

$$\begin{aligned}
\mu(\lambda) &= E(X|A) = \mu + \frac{\sigma^2}{P(|X| > a)}\left[\varphi_{(\mu,\sigma^2)}(a) - \varphi_{(\mu,\sigma^2)}(-a)\right] \\
\sigma^2(\lambda) &= Var(X|A) = \sigma^2 + \frac{\sigma^2}{P(|X| > a)}\left[(a - \mu)\varphi_{(\mu,\sigma^2)}(a) + (a + \mu)\varphi_{(\mu,\sigma^2)}(-a)\right] \\
&\quad - (\mu - \mu_A)^2.
\end{aligned}$$

Further, the expected value for one-sided conditions is given as

$$E(X \mid X > a) = \mu + \frac{\sigma^2}{P(X > a)}\left[\varphi_{(\mu,\sigma^2)}(a)\right] \tag{53}$$

$$E(X \mid X < -a) = \mu - \frac{\sigma^2}{P(X < -a)}\left[\varphi_{(\mu,\sigma^2)}(-a)\right]. \tag{54}$$

Proof For $X \sim N(\mu, \sigma^2)$ with density $\varphi_{(\mu,\sigma^2)}$, we have

$$\frac{d}{dx}\varphi_{(\mu,\sigma^2)}(x) = -\frac{x-\mu}{\sigma^2}\varphi_{(\mu,\sigma^2)}(x).$$

Now, for $A = \{X < -a \ \wedge \ X > a\}$, we use

$$\int_{-\infty}^{-a} \frac{d}{dx}\varphi_{(\mu,\sigma^2)}(x)dx = \varphi_{(\mu,\sigma^2)}(-a) \ \text{ and } \ \int_{a}^{\infty} \frac{d}{dx}\varphi_{(\mu,\sigma^2)}(x)dx = -\varphi_{(\mu,\sigma^2)}(a),$$

and obtain

$$\mathrm{E}\left(XI_A(X)\right) = \int_A x\varphi_{(\mu,\sigma^2)}(x)dx \ = \ \sigma^2\left[\varphi_{(\mu,\sigma^2)}(a) - \varphi_{(\mu,\sigma^2)}(-a)\right] + \mu P(A)$$

which gives

$$\mathrm{E}(X \mid |X| > a) = \mu + \frac{\sigma^2}{P(|X| > a)}\left[\varphi_{(\mu,\sigma^2)}(a) - \varphi_{(\mu,\sigma^2)}(-a)\right]. \tag{55}$$

Equations (53) and (54) follow analogously. Now, using

$$\frac{d^2}{dx}\varphi_{(\mu,\sigma^2)}(x) = \frac{1}{\sigma^4}\left((x-\mu)^2 - \sigma^2\right)\varphi_{(\mu,\sigma^2)}(x)$$

the variance of truncated normal distribution can be derived as

$$\mathrm{Var}(X \mid X \in A) = \frac{1}{P(A)}\int (x-\mu)^2\varphi_{(\mu,\sigma^2)}(x)I_A(x)dx - (\mu - \mu_A)^2 = V_1 - (\mu - \mu_A)^2.$$

In particular, for $A = \{|X| > a\}$, the integral part, V_1, follows as

$$\begin{aligned}
V_1 &= \frac{1}{P(A)}\left(\sigma^4\int_A \frac{d^2}{dx}\varphi_{(\mu,\sigma^2)}(x)dx + \sigma^2 P(A)\right) \\
&= \sigma^2 + \frac{\sigma^4}{P(|X| > a)}\left(\int_{-\infty}^{-a}\frac{d^2}{dx}\varphi_{(\mu,\sigma^2)}(x)dx + \int_a^{\infty}\frac{d^2}{dx}\varphi_{(\mu,\sigma^2)}(x)dx\right) \\
&= \sigma^2 + \frac{\sigma^4}{P(|X| > a)}\left(\frac{d}{dx}\varphi_{(\mu,\sigma^2)}(-a) - \frac{d}{dx}\varphi_{(\mu,\sigma^2)}(a)\right) \\
&= \sigma^2 + \frac{\sigma^2}{P(|X| > a)}\left[(a-\mu)\varphi_{(\mu,\sigma^2)}(a) + (a+\mu)\varphi_{(\mu,\sigma^2)}(-a)\right].
\end{aligned}$$

□

Proof of Theorem 1

Write $F(\boldsymbol{\beta}) = F_1(\boldsymbol{\beta}) + F_2(\boldsymbol{\beta})$, where F_1 is quadratic form and F_2 is $\mathbb{L}_1$ norm. For (a), the convexity of F follows from that of F_1 and F_2 [11, Proposition B.24(d), p. 732]. Now, $\partial F_1(\boldsymbol{\beta}) = -2\mathbf{X}^T\mathbf{W}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$ [7, Chap. 15]. For the sub-differential of $F_2(\boldsymbol{\beta})$, first note that [11, Proposition B.24(a), p. 732] $\partial F_2(\boldsymbol{\beta}) = \mathbf{A}^T\partial F_2(\boldsymbol{\beta})$. Then, with $\text{Dom}F_2(\boldsymbol{\beta}) \in \mathbb{R}^p$, $\partial F_2(\boldsymbol{\beta}) = \boldsymbol{\omega} \in \mathscr{L}(\mathbf{A}\boldsymbol{\beta})$, where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_p)^T$ is the sign function [12, p. 354] with $\omega_j = \text{sgn}(b_j)$ if $b_j \neq 0$ and $\omega_j \in [-1, 1]$ if $b_j = 0$. Combining the two components gives $\partial F(\boldsymbol{\beta})$ as in (14).

For (b), following the theory of convex analysis, the KKT conditions

$$-2\langle \mathbf{X}_{(j)}, \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\rangle_{\mathbf{W}} + \lambda\omega_j, \quad j = 1, \ldots, p,$$

provide necessary and sufficient condition for the solution of $F(\boldsymbol{\beta})$, where $\omega_j = \omega_j(\boldsymbol{\beta})$ is defined above and $\langle\cdot,\cdot\rangle$ denotes the inner product. Slightly rearranged, the solution can be written as in (16). □

Proof of Theorem 2

The proof holds easily for both $\widehat{\boldsymbol{\alpha}}$ to $\widehat{\boldsymbol{\beta}}$ or vice versa where the later is easier. In this case, first the proof for $\widehat{\boldsymbol{\beta}}$ follows from that of Theorem 1(a) using $\mathbf{A} = \mathbf{V}^T$. This gives $\mathbf{X}^T\mathbf{W}\mathbf{X}\boldsymbol{\beta} = \mathbf{X}^T\mathbf{W}\mathbf{Y} - (\lambda/2)\mathbf{V}\boldsymbol{\omega}$. By the SVD in (5), $\mathbf{X} = \mathbf{U}\mathbf{L}\mathbf{V}^T$, the result for $\widehat{\boldsymbol{\alpha}}$ follows for $\mathbf{X} = \mathbf{Z}\mathbf{V}^T$ with $\mathbf{Z} = \mathbf{U}\mathbf{L}$ and $\mathbf{V}^T\boldsymbol{\beta} = \boldsymbol{\alpha}$. □

Proof of Corollary 1

For given $\mathbf{W}$ and $\mathbf{A}$, with $\mathbf{V}$ orthonormal, Theorem 1(b) gives $\mathbf{Z}^T\mathbf{Z}\widehat{\alpha} = \mathbf{Z}^T\mathbf{Y} - \frac{\lambda}{2}\widehat{\boldsymbol{\omega}}$. The result then follows from (20). □

Proof of Theorem 4

For (31), consider $E(\widehat{\alpha}_j) = E\left(\widehat{\alpha}_j \mid \widehat{\alpha}_j \neq 0\right)(1 - p_{0j})$ with (see 22)

$$E\left(\widehat{\alpha}_j \mid \widehat{\alpha}_j \neq 0\right) = E\left(z_{(j)} - a_j(\lambda)\widehat{\omega}_j \mid \widehat{\alpha}_j \neq 0\right), \quad j = 1, \ldots, k.$$

The conditional distribution of $\widehat{\omega}_j$ is a two-point distribution with support $\{-1, 1\}$, so that, for $j = 1, \ldots, k$,

$$P\left(\widehat{\omega}_j=1 \mid \widehat{\alpha}_j \neq 0\right)=\frac{\mathrm{P}\left(\widehat{\alpha}_j>0\right)}{\mathrm{P}\left(\widehat{\alpha}_j \neq 0\right)}=\frac{p_{+j}}{1-p_{0j}} \Rightarrow E\left(\widehat{\omega}_j \mid \widehat{\alpha}_j \neq 0\right)=\frac{p_{+j}-p_{-j}}{1-p_{0j}},$$

which gives (31), where p_{+j}, p_{-j} and p_{0j} are as defined in (27). For (32), we note that

$$\begin{aligned}\operatorname{Var}(\widehat{\alpha}_j) &= \operatorname{Var}\left(\widehat{\alpha}_j \mid \widehat{\alpha}_j \neq 0\right)(1-p_{0j})+\operatorname{Var}\left(\widehat{\alpha}_j \mid \widehat{\alpha}_j=0\right) p_{0j} \\ &= \operatorname{Var}\left(\widehat{\alpha}_j \mid \widehat{\alpha}_j \neq 0\right)(1-p_{0j})\end{aligned}$$

where

$$\operatorname{Var}\left(\widehat{\alpha}_j \mid \widehat{\alpha}_j \neq 0\right)=\operatorname{Var}\left(z_{(j)}-a_j(\lambda)\widehat{\omega}_j \mid \widehat{\alpha}_j \neq 0\right)=\mathrm{V}_1+\mathrm{V}_2-2\mathrm{C}_{12}$$

with

$$\begin{aligned}\mathrm{V}_1 &= \operatorname{Var}(z_{(j)} \mid \widehat{\alpha}_j \neq 0)=\operatorname{Var}(z_{(j)} \mid \left|z_{(j)}\right|>a_j(\lambda)) \\ \mathrm{V}_2 &= a_j(\lambda)^2 \operatorname{Var}(\widehat{\omega}_j \mid \widehat{\alpha}_j \neq 0) \\ \mathrm{C}_{12} &= a_j(\lambda)\operatorname{Cov}(z_{(j)}, \widehat{\omega}_j \mid \widehat{\alpha}_j \neq 0).\end{aligned}$$

Applying Lemma 1 with $p=p_{+j}/(1-p_{0j})$,

$$\mathrm{V}_2=4a_j(\lambda)^2\frac{p_{+j}p_{-j}}{(1-p_{0j})^2} \text{ and } \mathrm{C}_{12}=2a_j(\lambda)\frac{p_{+j}p_{-j}}{(1-p_{0j})^2}e_j$$

where $e_j=\mathrm{E}(z_{(j)} \mid \widehat{\alpha}_j \neq 0, \omega_j=1)-\mathrm{E}(z_{(j)} \mid \widehat{\alpha}_j \neq 0, \omega_j=-1)=e_{j1}-e_{j2}$. Since $\widehat{\alpha}_j \neq 0, \omega_j=1 \Leftrightarrow \widehat{\alpha}_j>0$ and $\widehat{\alpha}_j \neq 0, \omega_j=-1 \Leftrightarrow \widehat{\alpha}_j<0$, therefore

$$e_{j1}=\mathrm{E}\left(z_{(j)} \mid \widehat{\alpha}_j>0\right)=\mathrm{E}\left(z_{(j)} \mid z_{(j)}>a_j(\lambda)\right)$$

and similarly $e_{j2}=\mathrm{E}(z_{(j)} \mid \widehat{\alpha}_j<0)=\mathrm{E}(z_{(j)} \mid z_{(j)}<-a_j(\lambda))$, and e_j is as given in (26). This gives C_{12}. Combining the results, we get (32) for $j=1,\ldots,k$. □

Proof of Corollary 2

We specialize the proof of Theorem 4 to the normal case. First, using truncated moments from Theorem 7, we get

$$E\left(z_{(j)} \mid \widehat{\alpha}_j \neq 0\right)=E\left(z_{(j)} \mid |z_{(j)}|>a_j(\lambda)\right)=\alpha_j+\frac{1}{\lambda_j^2(1-p_{0j})}d_j=\mu_j(\lambda),$$

with d_j as in (39). Substitution in (31) gives (35). For variance, consider V_1, V_2, C_{12} in the proof of Theorem 4 and note, using again Theorem 7, that

$$V_1 = \text{Var}\left(z_{(j)} \mid \left|z_{(j)}\right| > a_j(\lambda)\right) = \frac{1}{\lambda_j^2}\left[1 + \frac{1}{1-p_{0j}}v_j - \frac{1}{\lambda_j^2(1-p_{0j})^2}d_j^2\right],$$

with v_j as in (40). V_2 and C_{12} follow similarly. Substitution in (25) gives (36). Finally, $\text{MSE}(\widehat{\boldsymbol{\alpha}}) = \text{E}\|\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\|^2 = \sum_{j=1}^k \text{E}(\widehat{\alpha}_j - \alpha_j)^2 + \sum_{j=k+1}^p \alpha_j^2$ with $\text{E}(\widehat{\alpha}_j - \alpha_j)^2$ simplifying to the expression given in Theorem 4. □

Proof of Theorem 6

Given $z_{(j)} \sim N(\alpha_j, \lambda_j^{-2})$ and d_j, v_j, e_j as in (39)–(41). For simplicity, we may omit index j and write α. Further, we shall assume $\lambda_j = 1 \,\forall\, j$ and refer to it as the special case below. In this case, $z_{(j)} \sim N(\alpha, 1)$, $a_j(\lambda) = \lambda/2$, $d_j = d(\lambda, 1, \alpha) = \varphi(c_2) - \varphi(c_1)$, $e_j = e(\lambda, 1, \alpha) = \varphi(c_2)/p_+ - \varphi(c_1)/p_-$, $v_j = v(\lambda, 1, \alpha) = -c_2\varphi(c_2) + c_1\varphi(c_1)$, where $c_1 = \alpha + \lambda/2$, $c_2 = \alpha - \lambda/2$. Further, $p_0(\lambda) = \Phi(c_1) - \Phi(c_2)$, $p_+(\lambda) = \Phi(c_2)$, $p_-(\lambda) = 1 - \Phi(c_1)$ and $d(\lambda) = \varphi(c_2) - \varphi(c_1)$. Here, $\varphi(\cdot)$ and $\Phi(\cdot)$ denote, respectively, the density and distribution functions of the standard normal distribution.

To begin with the proofs, note that

$$p_0(-\alpha) = \Phi(-c_1) - \Phi(-c_2) = p_0(\alpha), \quad p_+(-\alpha) = 1 - \Phi(c_2) = p_-(\alpha);$$

similarly, $p_-(-\alpha) = p_+(\alpha)$ and $d(\alpha) = -d(-\alpha)$. This proves I. Now consider (37). For II, we first write

$$\text{bias}\left(\alpha_j\right) = \alpha_j p_{0j} + a_j(\lambda)\left(p_{+,j} - p_{-j}\right) - d_j/\lambda_j^2, \quad j = 1, \ldots, k.$$

Considering the expressions in I now as functions of α, we have

$$\frac{d[d(\alpha)]}{d\alpha} = -c_1\varphi(c_1) + c_2\varphi(c_2), \quad \frac{d[p_0(\alpha)]}{d\alpha} = \varphi(c_2) - \varphi(c_1) = -d(\alpha),$$

similarly $d[p_-(\alpha)]/d\alpha = -\varphi(c_1)$, so that $d[\text{bias}(\alpha)]/d\alpha = p_0(\alpha) > 0$. To prove III, the bias, considered as function of λ, reduces to

$$\text{bias}(\lambda) = \alpha p_0(\lambda) + \lambda\left[p_+(\lambda) - p_-(\lambda)\right]/2 - d(\lambda)$$

which, using $p_0 + p_+ + p_- = 1$, can be rewritten as

$$\text{bias}(\lambda) = c_2 p_0(\lambda) + \lambda/2 - \left[d(\lambda) + \lambda p_-(\lambda)\right].$$

With $d[\Phi(c_1)]/d\lambda = \varphi(c_1)/2$, $d[\varphi(c_1)]/d\lambda = -c_1\varphi(c_1)/2$, similarly for c_2, we get

$$\frac{d}{d\lambda}p_0(\lambda) = \frac{1}{2}\varphi(c_1) + \frac{1}{2}\varphi(c_2), \quad \frac{d}{d\lambda}p_-(\lambda) = -\frac{1}{2}\varphi(c_1),$$

so that $d[d(\lambda) + \lambda p_-(\lambda)]/d\lambda = c_2[\varphi(c_2) + \varphi(c_1)]/2 + p_-(\lambda)$. It eventually follows that $d[\text{bias}(\lambda])/d\lambda = (p_+ - p_-)/2 > 0$. Note that, with $\alpha > 0$ w.o.l.o.g., $p_+ > p_-$.

For IV, consider the variance as function of λ, i.e.,

$$\text{Var}_0(\lambda) = 1 - p_0(\lambda) + \lambda\varphi\left(\frac{\lambda}{2}\right) + \frac{\lambda^2}{2}p_+(\lambda).$$

Then, using the computations above, it follows that $d[p_0(\lambda)]/d\lambda = \varphi(\lambda/2)$ and $d[p_+(\lambda)]/d\lambda = -\frac{1}{2}\varphi(\lambda/2)$, so that

$$\begin{aligned}\frac{d}{d\lambda}\text{Var}_0(\lambda) &= -\frac{\lambda^2}{4}\varphi\left(\frac{\lambda}{2}\right) + \lambda\left(1 - \Phi\left(\frac{\lambda}{2}\right)\right) - \frac{\lambda^2}{4}\varphi\left(\frac{\lambda}{2}\right)\\ &= -\frac{\lambda^2}{2}\varphi\left(\frac{\lambda}{2}\right) + \lambda\Phi\left(-\frac{\lambda}{2}\right).\end{aligned}$$

It can then be verified that, $d[\text{Var}_0(\lambda)]/d\lambda \gtreqless 0$ for $\lambda \lesseqgtr 1.5036$, respectively.

Finally V, where we evaluate the MSE for $\lambda \to 0$ and $\lambda \to \infty$. For the first case, it can be verified that both $p_-(\lambda)$ and $p_+(\lambda)$ converge to 1/2, making $p_0(\lambda) \to 0$. Further, $\phi(c_1) = \phi(c_2)$ and $d(\lambda)$, $e(\lambda)$, $\nu(\lambda)$ and $a_j(\lambda)$ all vanish, so that bias$(\widehat{\alpha}_j) \to \alpha_j^2$ or bias$(\widehat{\boldsymbol{\alpha}}) \to \sum_{j=k+1}^{p}\alpha_j^2$, which vanish for $\alpha_j = 0$. Using the same set up for variance, we get var$(\widehat{\alpha}_j) \to 1/\lambda_j^2$ or var$(\widehat{\boldsymbol{\alpha}}) \to \sum_{j=1}^{k} 1/\lambda_j^2$ which reduces to k for $\lambda_j = 1 \;\forall\; j$. Combined, this gives (50).

The computations are relatively more involved for the case $\lambda \to \infty$, particularly for variance. First, $a_j(\lambda) \to \infty$ and $\phi(\infty) = 0$. Moreover, $d(\lambda)$ and $\nu(\lambda)$ vanish whereas $e(\lambda)$, by an application of L'Hospital rule, simplifies to $-2\alpha_j$ and thus vanish if $\alpha = 0$. Then, the first part of bias converges to α_j and the rest vanish so that bias$(\widehat{\boldsymbol{\alpha}}) \to \sum_{j=k+1}^{p}\alpha_j^2$. For variance, we particularly need to take care of the last two components. Using Bernstein's inequality [13], $\text{P}(|X - \mu| \geq \tau) \leq 2\exp(-\tau^2/2\sigma^2)$ for $X \sim N(\mu, \sigma^2)$, it follows, after some simplification, that

$$\frac{d^2}{1 - p_{0j}} \leq \exp\left(\frac{(a(\lambda) + |\alpha|)^2}{2/\lambda^2}\right)\varphi_{(\alpha,\lambda^2-2)}(a(\lambda))^2,$$

or $\exp(\{\lambda/2 + |\alpha_j|^2\}/2)\varphi_{(\alpha,1)}(\lambda/2)^2$ for the special case, and thus vanishes. The last component likewise vanishes using $e(\lambda)$ above. Combining all components, it implies that the variance vanishes for $\lambda \to \infty$. This gives (51). □

References

1. Tibshirani, R.J.: Regression shrinkage and selection via the LASSO. JRSS B **58**(1), 267–288 (1996)
2. Gruber, M.H.J.: Improving Efficiency by Shrinkage: The James–Stein and Ridge Regression Estimators. Marcel Dekker Inc, New York (1998)
3. Hoerl, A.E., Kennard, R.W.: Ridge Regression: Applications to Nonorthogonal Problems. Technometrics **12**(1), 55–67 and 69–82 (1970)
4. Bühlmann, P., van de Geer, S.: Statistics for High-Dimensional Data: Models, Theory & Applications. Springer, New York (2011)
5. Hastie, T., Tibshirani, R.J., Waineright, M.: Statistical Learning with Sparsity: The LASSO and Generalizations. CRC Press, Boca Raton (2015)
6. Gentle, J.E.: Matrix Algebra: Theory, Computations & Applications in Statistics. Springer, New York (2007)
7. Harville, D.A.: Matrix Anlgebra from a Statistician's Perspective. Springer, NY (1997)
8. Dicker, L.H.: Ridge regression and asymptotic minimax estimation over spheres of growing dimension. Bernoulli **22**(1), 1–37 (2016)
9. Gruber, Marvin H.J.: Liu and ridge estimator: a comparison. Commun. Stat.: Theory Methods **41**, 3739–3749 (2012)
10. Johnson, N.L., Kotz, S., Balakrishnan, N.: Continuous Univariate Distributions, vol. 1. Wiley, New York (1994)
11. Bertsekas, D.P.: Nonlinear programming, 2nd edn. Athena Scientific, MA (1999)
12. Lange, K.: Optimization, 2nd edn. Springer, NY (2013)
13. Bickel, P.J., Doksum, K.A.: Mathematical Statistics: Basic Ideas and Selected Topics. Prentice Hall, NJ (2001)