Abdelaati Daouia
Anne Ruiz-Gazen *Editors*

Advances in Contemporary Statistics and Econometrics

Festschrift in Honor of Christine Thomas-Agnan

Springer

Advances in Contemporary Statistics and Econometrics

Published with permission from Clara and Didier Gazen

Abdelaati Daouia · Anne Ruiz-Gazen
Editors

Advances in Contemporary Statistics and Econometrics

Festschrift in Honor of Christine Thomas-Agnan

Springer

Editors
Abdelaati Daouia
Toulouse School of Economics
University of Toulouse Capitole
Toulouse, France

Anne Ruiz-Gazen
Toulouse School of Economics
University of Toulouse Capitole
Toulouse, France

ISBN 978-3-030-73248-6 ISBN 978-3-030-73249-3 (eBook)
https://doi.org/10.1007/978-3-030-73249-3

Mathematics Subject Classification: 62GXX, 62HXX, 62JXX, 62P20

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

Foreword

I first met Christine in the spring of 2002 when I visited GREMAQ at Université Toulouse I, at her invitation to teach a short-course in spatial econometrics for graduate students. June 14 of that year was also the first *Workshop on Spatial Econometrics and Statistics*, organized by Christine, with Noel Cressie giving a keynote presentation. At that workshop I met Cem Ertur, Julie Le Gallo, and Catherine Baumont from Université de Bourgogne in Dijon who hosted the next workshop in May 2003. I also met Olivier Parent, a graduate student at the time, who drove from Strasbourg to Toulouse to attend the conference. In the years following these first two workshops, I had the good fortune to collaborate with many of the French researchers I met as well as their graduate students and to attend seven more of the workshops. Over the years, these workshops have been held in Avignon, Besançon, Dijon, Grenoble, Orleans, Paris, Strasbourg, and Toulon, and have attracted an international audience and several invited speakers from around the world. The success of this workshop has continued with the 19th workshop originally scheduled for May 2020 re-scheduled to May 2021 in Nantes due to the Covid-19 outbreak. This is one important legacy of Christine for those working in the areas of spatial statistics and econometrics.

This book contains five parts that reflect research areas in which Christine has worked over the years. These include nonparametric statistics and econometrics, quantiles and expectiles, spatial statistics and econometrics, compositional data analysis, and tools for empirical studies in economics and applications. So, spatial statistics and econometrics reflects only one area of Christine's past research efforts, but the continuing success of the workshop is a wonderful example of her ability to bring together researchers and promote collaboration.

In terms of collaboration, Christine has worked with nearly 50 co-authors during the last 25 years, on publications appearing in prestigious journals such as *Journal of the American Statistical Association*, *Annals of Statistics*, *Econometric Theory*, *Statistical Papers*, *Journal of Regional Science*, *Numerical Algorithms*, *Statistics & Probability Letters*, *Statistical Methodology*, *Journal of Nonparametric Statistics*, and *Computational Statistics*. Her 2004 book *Reproducing kernel Hilbert spaces in probability and statistics* co-authored with Alain Berlinet has received a great deal of attention in the literature, as has her *Econometric Theory* article: Nonparametric frontier estimation: a conditional quantile-based approach, co-authored with Yves Aragon and Abdelaati Daouia.

This volume contains numerous contributions, some by those who have collaborated with Christine over the years, and we can all learn from these works. I have collaborated with Christine and found her to be a fabulous person to work with and have benefited from her scholarly expertise and insights. Christine is one of only a handful of scholars whose interests span very theoretical statistical issues as well as applied research that aims at tackling real-world problems. The contributions in this volume reflect that broad range of interests, so there is something for everyone to enjoy. Christine's past work and her ability to promote collaboration among researchers have been an inspiration to us all. Let us hope she continues her research and collaborative efforts for many years to come.

November 2020

James P. LeSage
Fields Endowed Chair
Texas State University
San Marcos, USA

Preface

Christine Thomas-Agnan became Senior Lecturer at Toulouse Capitole University (UTC) in 1988 and Professor in 1994, after completing a doctoral thesis at the University of California, Los Angeles, and a PRAG teaching position at Toulouse Jean Jaurès University. She founded and chaired the group STATISTIQUE—UT1 from 1994, when the faculty of Economics of UTC moved to the building of "Manufacture des Tabacs de Toulouse". Currently, she heads the mathematics department at UTC.

Over her long and brilliant academic career, Christine Thomas-Agnan has worked on a variety of topics in mathematical and applied statistics, including nonparametric and semi-parametric inference, spatial statistics and econometrics, compositional data analysis, market share regression models, and political economics statistical models. Her work in these areas has found applications in a broad variety of fields, including efficiency measurements of French postal services and Spanish electricity distributors, optimal location of a new fire station in the surroundings of Toulouse, explaining the patterns of regional unemployment and doctors' prescribing in the Midi-Pyrénées region, testing spatial dependence in air passenger flows, assessing the relations between socioeconomic factors and nutritional diet in Vietnam, and understanding the impact of the composition of media investments on automobile sales in the French automobile market, to cite a few. She has published 5 books and over 50 refereed works in top academic journals. She has been Chief Editor of the Journal CSBIGS (Case Studies in Business, Industry and Government) since 2015, and Member of the publications committee of the French Statistical Society. She has also supervised 12 Ph.D. students and 5 Habilitation degrees (HDR).

Christine Thomas-Agnan is not only a gifted and inspirational researcher and teacher but also a hard-working colleague with a fruitful and curious mind. She has boundless and communicative energy that she puts at the service of the University, her colleagues, co-authors, and students at all levels, especially her Ph.D. students. Her enthusiasm and open-mindedness are greatly appreciated by all. Working with her is an absolute pleasure for us, researchers and teachers in the statistics group, and more generally in the mathematics department, as she facilitated a high-level stimulating environment while maintaining a friendly and inviting demeanor that makes us feel like family.

The task of editing this volume was remarkably easy as the colleagues contacted were so enthusiastic about contributing to this Festschrift by writing and/or editing a research article in Christine's honor. She had dozens of collaborators on an extraordinary variety of research topics. As evidenced by the many tributes in this volume, all colleagues who have had the chance to work with Christine praise her human and scientific qualities.

The 35 articles in this volume are at the frontier of contemporary research in the fields of statistics and econometrics. They testify to Christine's numerous contributions in these fields at both theoretical and applied levels. Christine was first trained as a specialist in reproducing kernel Hilbert space theory and its use in statistical applications. The results she has established since her Ph.D. thesis were published in 2004 in a Springer book jointly with Alain Berlinet. In 1987, she started to explore nonparametric regression by elegantly using spline and kernel smoothing. Then, in 1993, she oriented her research toward functional estimation under form constraints. In the meantime, her intense work on nonparametric and semi-parametric modeling led her to the active fields of quantile/expectile regression and dimension reduction for multivariate response data. In 2002, she began to orient her research toward spatial statistics and econometrics through collaborations she initiated with James P. LeSage and Noel Cressie at the first spatial econometrics workshop she organized at UTC. By adopting the mathematical rigor of statistics and benefiting from the subtlety of econometrics, Christine has first generalized existing models to take into account spatial autocorrelation, and investigated Monte Carlo estimation of Markovian Gaussian fields, before moving to spatial point processes and their use to deal with spatial homogeneity tests, cluster detection, and optimal location-allocation problems. Her efforts have also focused on combining nonparametric methods with spatial statistics to estimate, for instance, autocovariance functions not only of processes but also of random fields, and to study the implications on kriging. Christine's attention was also directed toward the area of frontier and efficiency analysis in production econometrics, with her influential 2005 *Econometric Theory* paper in this literature. From 2011 to 2016, she has been the principal investigator of the interdisciplinary *ModULand* project on the modeling of land use, a prestigious research grant of the French National Research Agency. More recently, she has become interested in compositional data analysis and market share regression models with a particular attention to measuring the impact of covariates in spatial and compositional models. Her recent research allows her to investigate new areas while integrating various interdisciplinary components of her previous research.

Christine's impressive research record should not, however, hide her immense investment in education and services to the community and students. For more than 30 years, she has been heavily involved in the Master program of Econometrics and Statistics at the faculty of Economics of UTC and more recently at Toulouse School of Economics. Among other things, she created the statistical consultancy course of the Master 2 in Statistics and Econometrics more than 20 years ago. This course allows our students to develop their ability to confront concrete statistical problems, posed by companies, under reassuring university supervision. Students were given an invaluable opportunity to experience concrete and exciting projects.

As she nears retirement, one might think that Christine would have less energy or desire to invest in new areas or experience new things in her work, but nothing could be further from the truth. In addition to assuming responsibility for the Master 1 in Econometrics and Statistics at the Toulouse School of Economics, she has accepted the direction of the mathematics department, which she manages with great initiative and tact. She has also very recently accepted to supervise new doctoral students, namely Lukas Dargel and Thibault Laurent, in two stimulating research programs with applications in social sciences.

We would like to thank Christine Thomas-Agnan for being such an inspiring figure in our professional and personal lives. We join our colleague James P. LeSage in the hope that she will continue her excellent work for many years to come.

The five parts of this volume correspond to the topics that Christine has contributed much to. The contributions collected in each section answer important questions that reflect varied theoretical and/or applied interests of their authors. They provide nice examples of the new research ideas that are currently being developed. We expect that everyone will find something interesting in this rich collection of papers.

The first part contains seven papers related to the active area of nonparametric statistics and econometrics. Fadoua Balabdaoui and Piet Groeneboom elucidate the open question of whether a profile least squares estimator in the monotone single index model is $\sqrt{n}$ convergent and asymptotically normal. Gérard Biau and Benoît Cadre present a general framework for studying two widespread gradient boosting algorithms from the perspective of functional optimization, and address the less-discussed problem of their convergence as the number of iterations tends to infinity. Sandrine Casanova and Eve Leconte introduce a novel nonparametric model-based estimator for the conditional distribution function of a right censored response, which is superior to its most known competitors in small domains. Eric Gautier suggests endogenous selection models, which allow for instrument nonmonotonicity and are based on nonparametric random coefficient indices. Camelia Goga gives a review of applications of B-spline regression in a survey sampling framework and design-based approach, including new properties of the (un)penalized estimators, and their improved consistency rates. Hadrien Lorenzo and Jérôme Saracco propose three computational devices to detect outliers in a single index regression model, when conducting sliced inverse regression along with kernel smoothing of the link function. Jan Meis and Enno Mammen revisit the uncoupled isotonic regression problem by improving the rate of convergence of the so-called minimum Wasserstein deconvolution estimator, for L_p-risks and for error distributions supported on a finite set of points.

The second part also contains seven contributions that are dedicated to the topic of (un)conditional quantiles and expectiles. The class of expectiles corresponds to a least squares analogue of quantiles. Cécile Adam and Irène Gijbels study multivariate partially linear expectile regression in which the nonlinear part is fitted using a local polynomial approach, along with an optimal choice of the bandwidth parameter. Delphine Blanke and Denis Bosq prove that, for estimating univariate quantiles, the reciprocal of the piecewise linear interpolation at the midpoints of a sample distribution function strictly improves the MISE of the usual sample quantile function. Axel

Bücher, Anouar El Ghouch, and Ingrid Van Keilegom propose a valid local linear smoothing approach to iteratively estimate a semi-parametric single-index model for conditional quantiles with right-censored data. Stéphane Girard, Gilles Stupfler, and Antoine Usseglio-Carleve construct kernel estimators of extreme regression L_p-quantiles, which encompass both families of expectiles and standard quantiles, and develop their asymptotic theory for heavy-tailed conditional distributions. Bao Hoang Nguyen and Valentin Zelenyuk perform a robust frontier and efficiency analysis of public hospitals in Queensland, Australia, by estimating both individual and aggregate quantile-based efficiency scores. Davy Paindaveine and Joni Virta unravel the behavior of extreme d-dimensional spatial quantiles under minimal conditions, in a general setup for both population and sample multivariate distributions. Fabian Otto-Sobotka, Radoslava Mirkov, Benjamin Hofner, and Thomas Kneib use shape-constrained expectile regression in conjunction with a geoadditive model to provide deeper insights into the behavior of gas flow within transmission networks.

The third part concerns spatial statistics and econometrics with eight contributions. François Bachoc provides a review of the asymptotic theory for maximum likelihood estimation of covariance parameters for Gaussian processes, under increasing and fixed-domain asymptotics. Florent Bonneu and Lionel Cucala adapt spatial scan methods, borrowed from local cluster detection, to test for global similarity between two spatial point patterns. Hervé Cardot and Antonio Musolesi rely on the use of additive models and conditional mixtures and on random forests to estimate the variation along time of the spillover effects of spatial policies. Raja Chakir and Julie Le Gallo review the current state of the literature on studies which account for spatial autocorrelation in econometric land use models or in the environmental impacts of land use. Noel Cressie and Christopher Wikle develop a modern hierarchical statistical approach to modeling spatio-temporal data on regular or irregular spatial lattices. Van Huyen Do, Thibault Laurent, and Anne Vanhems implement widely used methods in the areal interpolation problem using R software, and provide practical guidelines to concrete questions such as spatial scales, types of target variable, and border incompatibility. Thibault Laurent and Paula Margaretic apply prediction of spatial econometric models for areal data to model regional unemployment rates taking into account local interactions. Mary Lai Salvaña and Marc Genton propose a new estimation methodology for nonstationary covariance models of the Lagrangian type, by modeling the second-order nonstationarity parameters via thin plate splines and estimating all the parameters via two-step maximum likelihood estimation.

The fourth part contains six papers on the area of compositional data analysis that Christine has also contributed to over the last years. Peter Filzmoser, Karel Hron, and Alessandra Menafoglio present and discuss a log-ratio approach to distributional modeling in a unifying framework for the discrete and the continuous distributional data based on the theory of Bayes spaces. Built on ideas from the spatial Durbin model, Tingting Huang, Gilbert Saporta, and Huiwen Wang propose and estimate a new compositional linear model for areal data by employing the orthonormal log-ratio transformation and maximum likelihood method. Wilfredo Maldonado, Juan José Egozcue, and Vera Pawlowsky-Glahn contribute to the modeling and compositional analysis of exchange rate matrices and the corresponding no-arbitrage matrices, by considering the Special Drawing Rights and by studying the relative exchange rate bubbles among the countries. Josep Antoni Martín-Fernández and Carles Barceló-Vidal revisit the basic concepts and properties of log ratios, log contrasts, and orthonormal coordinates for compositional data, and introduce a new approach that includes both the log-ratio orthonormal coordinates and an auxiliary variable carrying absolute information. Christoph Muehlmann, Kamila Fačevicová, Alžbčta Gardlo, Hana Janečková, and Klaus Nordhausen review some basic methods of independent component analysis and show how to apply such analysis to compositional data. Michel Simioni, Huong Thi Trinh, Tuyen Thi Thanh Huynh, and Thao-Vy Vuong explore the association between food sources and diet quality in Vietnam by making use of recent advances in compositional data analysis.

The seven contributions collected in the last part provide useful tools for empirical studies in economics and applied work. Bastien Bernela, Liliane Bonnal, and Pascal Favard untangle the empirical reality of the phenomenon of geographical mobility among students and young graduates in France. Christophe Bontemps and Valérie Orozco show how the research process, from data collection to paper publication, could efficiently be reorganized to improve and promote reproducible research. Olivier de Mouzon, Thibault Laurent, and Michel Le Breton explore and estimate the departure from the "One Man, One Vote" principle in the context of political representation and its consequences for distributive politics. They also provide several applications of the Lorenz curve and the Gini and Dauer-Kelsay indices to the measurement of malapportionment and disproportionality. Jonathan Haughton and Dominique Haughton recommend and illustrate the use of cartograms as an effective complement to the more-traditional choropleth maps for conveying spatially distributed statistical data. Jérôme Mariette, Madalina Olteanu, and Nathalie Vialaneix present kernel and dissimilarity methods to perform exploratory analysis in the presence of multiple sources of data or of multiple kernels describing different features of the data. Finally, Alban Thomas develops and applies a generalized method of particle nonlinear filtering to estimate a system of structural equations for agricultural crop yield functions, when unobserved productivity depends on water availability that is only partially observed.

Toulouse, France
January 2021

Abdelaati Daouia
Anne Ruiz-Gazen

Acknowledgements

All contributions are peer-reviewed, and we would like to thank all the referees very warmly for their very careful work.

Yasser Abbas
Yves Aragon
Cécile Adam
Denis Allard
François Bachoc
Fadoua Balabdaoui
Liliane Bel
Philippe Besse
Roger Bivand
Delphine Blanke
Liliane Bonnal
Florent Bonneu
Christophe Bontemps
Sandrine Casanova
Hervé Cardot
Raja Chakir
Noel Cressie
Lionel Cucala
Lukas Dargel
Philippe de Donder
Juan José Egozcue
Peter Filzmoser
Edith Gabriel
Sébastien Gadat
Serge Garcia
Eric Gautier
Marc G. Genton
Irène Gijbels
David Ginsbourger
Stéphane Girard
Antoine Godichon-Baggioni
Michel Goulard
Dominique Haughton
Xavier d'Haultfoeuille
David Haziza
Karel Hron
Thomas Kneib
Thibault Laurent
Pascal Lavergne
Eve Leconte
Julie Le Gallo
Erwan Le Pennec
Rik Lopuhaä
Enno Mammen
Paula Margaretic
Josep Antoni Martín-Fernández
Jorge Mateu
Joanna Morais
Christoph Muhlmann
Klaus Nordhausen
An Nguyen Huong
Madalina Olteanu
Jean Opsomer
Fabian Otto-Sobotka
Davy Paindaveine
Javier Palarea
Olivier Parent
Valentin Patilea
Vera Pawlowsky-Glahn
Gilbert Saporta
Jérôme Saracco
Michel Simioni
Gilles Stupfler
Alban Thomas
Anne Vanhems
Karine Van der Straten
Ingrid Van Keilegom
Nathalie Vialaneix
Joni Virta
Valentin Zelenyuk

Contents

Compositional Data Analysis

Tools for Empirical Studies in Economics and Social Sciences

Contributors

Cécile Adam Department of Mathematics, KU Leuven, Leuven (Heverlee), Belgium

François Bachoc Institut de Mathématiques, UMR; Université de Toulouse; CNRS, UPS IMT, Toulouse Cedex, France

Fadoua Balabdaoui Seminar für Statistik, ETH, Zürich, Switzerland

Carles Barceló-Vidal Department of IMAE, University of Girona, Girona, Spain

Bastien Bernela Université de Poitiers CRIEF, UFR de sciences économiques: rue Jean Carbonnier TSA, Poitiers cedex, France

Gérard Biau Sorbonne Université, CNRS, LPSM, Paris, France

Delphine Blanke Laboratoire de Mathématiques d'Avignon, LMA, Avignon Université, Avignon, France

Liliane Bonnal Université de Poitiers CRIEF and TSE, UFR de sciences économiques: rue Jean Carbonnier TSA, Poitiers cedex, France

Florent Bonneu Avignon Université, Avignon, France

Christophe Bontemps Toulouse School of Economics - INRAE, University of Toulouse Capitole, Toulouse, France

Denis Bosq Laboratoire de Probabilités, Statistique et Modélisation, LPSM, CNRS, Sorbonne Universités, Paris, France

Axel Bücher Mathematisches Institut, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany

Benoît Cadre Université Rennes, CNRS, IRMAR - UMR, Rennes, France

Hervé Cardot Université de Bourgogne, Institut de Mathématiques de Bourgogne, Dijon, France

Sandrine Casanova TSE-R, Université Toulouse Capitole, Toulouse cedex, France

Raja Chakir Université Paris-Saclay, INRAE, AgroParisTech, Economie Publique, France

Noel Cressie University of Wollongong, Wollongong, Australia

Lionel Cucala Université de Montpellier, Montpellier, France

Olivier de Mouzon Toulouse School of Economics, INRAE, University of Toulouse Capitole, Toulouse, France

Van Huyen Do Toulouse School of Economics, CNRS, University of Toulouse, Toulouse, France

Juan José Egozcue Technical University of Catalonia, Department of Civil and Environmental Engineering, Barcelona, Spain

Kamila Fačevicová Department of Mathematical Analysis and Applications of Mathematics, Palacký University Olomouc, Olomouc, Czech Republic

Pascal Favard Université de Tours IRJI, Department of Economics, Tours Cedex, France

Peter Filzmoser Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria

Alžběta Gardlo Department of Clinical Biochemistry, University Hospital Olomouc and Palacký University Olomouc, Olomouc, Czech Republic

Eric Gautier TSE, Université Toulouse Capitole, Esplanade de l'Université, Toulouse, France

Marc G. Genton King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

Anouar El Ghouch ISBA, UCLouvain, Louvain-la-Neuve, Belgium

Irène Gijbels Department of Mathematics and Leuven Statistics Research Center (LStat), Leuven (Heverlee), Belgium

Stéphane Girard University of Grenoble-Alpes, Inria, CNRS, Grenoble INP, LJK, Grenoble, France

Camelia Goga Laboratoire de Mathématiques de Besançon, Université de Bourgogne Franche-Comté, Besançon, France

Piet Groeneboom Delft University of Technology, DIAM, CD Delft, The Netherlands

Dominique Haughton Department of Mathematical Sciences, Bentley University, Waltham, MA, USA;
Université Paris I (SAMM), Paris, France;
Université Toulouse I (TSE-R), Toulouse, France

Jonathan Haughton Department of Economics, Suffolk University, Boston, MA, USA

Benjamin Hofner Section Biostatistics, Paul-Ehrlich-Institut, Langen, Germany

Karel Hron Department of Mathematical Analysis and Applications of Mathematics, Palacký University, Olomouc, Czech Republic

Tingting Huang School of Statistics, Capital University of Economics and Business, Beijing, China;
School of Economics and Management, Beihang University, Beijing, China;
Beijing Key Laboratory of Emergence Support Simulation Technologies for City Operations, Beijing, China

Tuyen Thi Thanh Huynh International Center for Tropical Agriculture (CIAT)—Asia Office, Hanoi, Vietnam

Hana Janečková Laboratory for Inherited Metabolic Disorders, Department of Clinical Biochemistry, University Hospital Olomouc and Palacký University Olomouc, Olomouc, Czech Republic

Thomas Kneib Department of Economics, Georg-August-Universität Göttingen, Göttingen, Germany

Thibault Laurent Toulouse School of Economics, CNRS, University of Toulouse Capitole, Toulouse, France

Michel Le Breton Institut Universitaire de France and Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France

Julie Le Gallo CESAER UMR, Agrosup Dijon, INRAE, Université de Bourgogne Franche-Comté, Dijon, France

Eve Leconte TSE-R, Université Toulouse Capitole, Toulouse cedex, France

Hadrien Lorenzo Inria BSO, Talence, France

Wilfredo L. Maldonado Faculty of Economics, Management and Accounting, University of São Paulo. Av. Professor Luciano Gualberto, São Paulo, Brazil

Enno Mammen Institute of Applied Mathematics, Heidelberg University, Heidelberg, Germany

Paula Margaretic University of San Andrés, Victoria, Argentina

Jérôme Mariette Université de Toulouse, INRAE, UR MIAT, Castanet-Tolosan, France

Josep Antoni Martín-Fernández Department of IMAE, University of Girona, Girona, Spain

Jan Meis Institute of Medical Biometry and Informatics, Heidelberg University, Heidelberg, Germany

Alessandra Menafoglio MOX, Department of Mathematics, Politecnico di Milano, Milano, Italy

Radoslava Mirkov Department of Mathematics, Humboldt University Berlin, Berlin, Germany

Christoph Muehlmann Institute of Statistics & Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria

Antonio Musolesi Department of Economics and Management and SEEDS, Ferrara University, Ferrara, Italy

Bao Hoang Nguyen School of Economics, University of Queensland, Brisbane, QLD, Australia

Klaus Nordhausen Institute of Statistics & Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria

Madalina Olteanu SAMM, Université Paris, Paris, France

Valérie Orozco Toulouse School of Economics - INRAE, University of Toulouse Capitole, Toulouse, France

Fabian Otto-Sobotka Division of Epidemiology and Biometry, Carl von Ossietzky University Oldenburg, Oldenburg, Germany

Davy Paindaveine Université libre de Bruxelles (ECARES and Department of Mathematics) and Université Toulouse Capitole (Toulouse School of Economics), Brussels, Belgium

Vera Pawlowsky-Glahn University of Girona, Department of Computer Science, Applied Mathematics and Statistics, Girona, Spain

Mary Lai O. Salvaña King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

Gilbert Saporta CNAM, Center for Studies and Research in Computer Science and Communication, Paris, France

Jérôme Saracco Inria BSO & ENSC Bordeaux INP, Talence, France

Michel Simioni MOISA, INRAE, University of Montpellier, Montpellier, France

Gilles Stupfler University of Rennes, Ensai, CNRS, CREST - UMR, Rennes, France

Alban Thomas Toulouse School of Economics-Research, INRAE, University of Toulouse, Toulouse, France

Huong Thi Trinh Department of Mathematics and Statistics, Thuongmai University, Hanoi, Vietnam

Antoine Usseglio-Carleve University of Grenoble-Alpes, Inria, CNRS, Grenoble INP, LJK, Grenoble, France

Ingrid Van Keilegom ORSTAT, KU Leuven, Leuven, Belgium

Anne Vanhems TBS Business School, Toulouse, France

Nathalie Vialaneix Université de Toulouse, INRAE, UR MIAT, Castanet-Tolosan, France

Joni Virta University of Turku (Department of Mathematics and Statistics) and Aalto University School of Science (Department of Mathematics and Systems Analysis), Turun yliopisto, Finland

Thao-Vy Vuong College of Agriculture and Life Sciences, Cornell University, Ithaca, USA

Huiwen Wang School of Economics and Management, Beihang University, Beijing, China;
Beijing Advanced Innovation Center for Big Data and Brain Computing, Beihang University, Beijing, China

Christopher K. Wikle University of Missouri, Columbia, USA

Valentin Zelenyuk School of Economics and Centre for Efficiency and Productivity Analysis, University of Queensland, Brisbane, QLD, Australia

Nonparametric Statistics and Econometrics

Profile Least Squares Estimators in the Monotone Single Index Model

Fadoua Balabdaoui and Piet Groeneboom

Abstract We consider least squares estimators of the finite regression parameter $\boldsymbol{\alpha}$ in the single index regression model $Y = \psi(\boldsymbol{\alpha}^T \boldsymbol{X}) + \varepsilon$, where $\boldsymbol{X}$ is a d-dimensional random vector, $\mathbb{E}(Y|\boldsymbol{X}) = \psi(\boldsymbol{\alpha}^T \boldsymbol{X})$, and ψ is a monotone. It has been suggested to estimate $\boldsymbol{\alpha}$ by a profile least squares estimator, minimizing $\sum_{i=1}^n (Y_i - \psi(\boldsymbol{\alpha}^T \boldsymbol{X}_i))^2$ over monotone ψ and $\boldsymbol{\alpha}$ on the boundary $\mathcal{S}_{d-1}$ of the unit ball. Although this suggestion has been around for a long time, it is still unknown whether the estimate is $\sqrt{n}$-convergent. We show that a profile least squares estimator, using the same pointwise least squares estimator for fixed $\boldsymbol{\alpha}$, but using a different global sum of squares, is $\sqrt{n}$-convergent and asymptotically normal. The difference between the corresponding loss functions is studied and also a comparison with other methods is given.

1 Introduction

The monotone single index model tries to predict a response from the linear combination of a finite number of parameters and a function linking this linear combination to the response via a monotone *link function* ψ_0 which is unknown. So, more formally, we have the model

$$Y = \psi_0(\boldsymbol{\alpha}_0^T \boldsymbol{X}) + \varepsilon,$$

where Y is a one-dimensional random variable, $\boldsymbol{X} = (X_1, \ldots, X_d)^T$ is a d-dimensional random vector with distribution function G, ψ_0 is monotone, and ε is a one-dimensional random variable such that $\mathbb{E}[\varepsilon|\boldsymbol{X}] = 0$ G almost surely. For identifiability, the regression parameter $\boldsymbol{\alpha}_0$ is a vector of norm $\|\boldsymbol{\alpha}_0\|_2 = 1$, where

F. Balabdaoui (✉)
Seminar für Statistik ETH, Zürich, Switzerland
e-mail: fadoua.balabdaoui@stat.math.ethz.ch

P. Groeneboom
Delft University of Technology, DIAM, Mekelweg 4, 2628 CD Delft, The Netherlands
e-mail: P.Groeneboom@tudelft.nl

A. Daouia and A. Ruiz-Gazen (eds.), *Advances in Contemporary Statistics and Econometrics*, https://doi.org/10.1007/978-3-030-73249-3_1

$\|\cdot\|_2$ denotes the Euclidean norm in $\mathbb{R}^d$, so $\boldsymbol{\alpha}_0 \in \mathcal{S}_{d-1}$, the unit $(d-1)$-dimensional sphere.

The ordinary profile least squares estimate of $\boldsymbol{\alpha}_0$ is an M-estimate in two senses: for fixed $\boldsymbol{\alpha}$, the least squares criterion

$$\psi \mapsto n^{-1} \sum_{i=1}^{n} \left\{Y_i - \psi(\boldsymbol{\alpha}^T \boldsymbol{X}_i)\right\}^2 \tag{1}$$

is minimized for all monotone functions ψ (either decreasing or increasing) which gives an $\boldsymbol{\alpha}$-dependent function $\hat{\psi}_{n,\boldsymbol{\alpha}}$, and the function

$$\boldsymbol{\alpha} \mapsto n^{-1} \sum_{i=1}^{n} \left\{Y_i - \hat{\psi}_{n,\boldsymbol{\alpha}}(\boldsymbol{\alpha}^T \boldsymbol{X}_i)\right\}^2 \tag{2}$$

is then minimized over $\boldsymbol{\alpha}$. This gives a profile least squares estimator $\hat{\boldsymbol{\alpha}}_n$ of $\boldsymbol{\alpha}_0$, which we will call LSE in the sequel. Although this estimate of $\boldsymbol{\alpha}_0$ has been known now for a very long time (more than 30 years probably), it is not known whether it is $\sqrt{n}$-convergent (under appropriate regularity conditions), let alone that we know its asymptotic distribution. Also, simulation studies are rather inconclusive. For example, it is conjectured in Tanaka (2008) on the basis of simulations that the rate of convergence of $\hat{\boldsymbol{\alpha}}_n$ is $n^{9/20}$. Other simulation studies, presented in Balabdaoui et al. (2019a), are also inconclusive. In that paper, it was also proved that an ordinary least squares estimator (which ignores that the link function could be non-linear) is $\sqrt{n}$-convergent and asymptotically normal under elliptic symmetry of the distribution of the covariate $\boldsymbol{X}$. Another linear least squares estimator of this type, where the restriction on $\boldsymbol{\alpha}$ is $\boldsymbol{\alpha}^T \boldsymbol{S}_n \boldsymbol{\alpha} = 1$, S_n is the usual estimate of the covariance matrix of the covariates, and a renormalization at the end is not needed (as it is in the just mentioned linear least squares estimator) was studied in Balabdaoui et al. (2019b) and was shown to have similar behavior. If this suggests that the profile LSE should also be $\sqrt{n}$-consistent, the extended simulation study in Balabdaoui et al. (2019b) shows that it is possible to find other estimates which exhibit better performance in these circumstances.

An alternative way to estimate the regression vector is to minimize the criterion

$$\boldsymbol{\alpha} \mapsto \left\| n^{-1} \sum_{i=1}^{n} \left\{Y_i - \hat{\psi}_{n,\boldsymbol{\alpha}}(\boldsymbol{\alpha}^T \boldsymbol{X}_i)\right\} \boldsymbol{X}_i \right\|^2 \tag{3}$$

over $\boldsymbol{\alpha} \in \mathcal{S}_{d-1}$, where $\|\cdot\|$ is the Euclidean norm. Note that this is the sum of d squares. The rational behind minimizing (3) is the fact that the true index vector, $\boldsymbol{\alpha}_0$, satisfies the (population) score equation

$$\mathbb{E}\left\{(Y - \psi_0(\boldsymbol{\alpha}_0^T \boldsymbol{X}))\boldsymbol{X}\theta(\boldsymbol{\alpha}_0^T \boldsymbol{X})\right\} = \boldsymbol{0}, \tag{4}$$

where θ is any measurable and bounded function. This clearly follows from the iterative law of expectations and the fact that $\mathbb{E}\{Y|\boldsymbol{\alpha}_0^T\boldsymbol{X}\} = \psi_0(\boldsymbol{\alpha}_0^T\boldsymbol{X})$. If the function θ is taken to be the constant 1, then the goal is to find the minimizer of the Euclidean norm of the empirical counterpart of the above score equation, after replacing the unknown link function, ψ_0, by its estimator $\hat{\psi}_{n,\boldsymbol{\alpha}}$.

We prove in Sect. 3 that this minimization procedure leads to a $\sqrt{n}$-consistent and asymptotically normal estimator, which is a more precise and informative result compared to what we know now about the LSE.. Using the well-known properties of isotonic estimators, it is easily seen that the function (3) is piecewise constant as a function of $\boldsymbol{\alpha}$, with finitely many values, so the minimum exists and is equal to the infimum over $\boldsymbol{\alpha} \in \mathcal{S}_{d-1}$. Notice that this estimator does not use any tuning parameters, just like the LSE.

In Balabdaoui et al. (2019b), a similar Simple Score Estimator (SSE) $\hat{\boldsymbol{\alpha}}_n$ was defined as a point $\boldsymbol{\alpha} \in \mathcal{S}_{d-1}$ where all components of the function

$$\boldsymbol{\alpha} \mapsto n^{-1}\sum_{i=1}^{n}\left\{Y_i - \hat{\psi}_{n,\boldsymbol{\alpha}}(\boldsymbol{\alpha}^T\boldsymbol{X}_i)\right\}\boldsymbol{X}_i$$

cross zero. If the criterion function were continuous in $\boldsymbol{\alpha}$, this estimator would have been the same as the least squares estimator, minimizing (3), with a minimum equal to zero, but in the present case we cannot assume this because of the discontinuities of the criterion function.

The definition of an estimator as a crossing of the d-dimensional vector $\mathbf{0}$ makes it necessary to prove the existence of such an estimator, which we found to be a rather non-trivial task. Defining our estimator directly as the minimizer of (3), so as a least squares estimator, relieves us from the duty to prove its existence. Since our estimator has the same limit distribution as the SSE, we refer to it here under the same name.

A fundamental function in our treatment is the function $\psi_{\boldsymbol{\alpha}}$, defined as follows.

Definition 1 Let $\mathcal{S}_{d-1}$ denote again the boundary of the unit ball in $\mathbb{R}^d$. Then, for each $\boldsymbol{\alpha} \in \mathcal{S}_{d-1}$, the function $\psi_{\boldsymbol{\alpha}} : \mathbb{R} \to \mathbb{R}$ is defined as the nondecreasing function which minimizes

$$\psi \mapsto \mathbb{E}\{Y - \psi(\boldsymbol{\alpha}^T\boldsymbol{X})\}^2$$

over all nondecreasing functions $\psi : \mathbb{R} \to \mathbb{R}$. The existence and uniqueness of the function $\psi_{\boldsymbol{\alpha}}$ follows, for example, from the results in Landers and Rogge (1981).

The function $\psi_{\boldsymbol{\alpha}}$ coincides in a neighborhood of $\boldsymbol{\alpha}_0$ with the ordinary conditional expectation function $\tilde{\psi}_{\boldsymbol{\alpha}}$

$$\tilde{\psi}_{\boldsymbol{\alpha}}(u) = \mathbb{E}\left\{\psi_0(\boldsymbol{\alpha}_0^T\boldsymbol{X})|\boldsymbol{\alpha}^T\boldsymbol{X} = u\right\}, \qquad u \in \mathbb{R}; \tag{5}$$

see Balabdaoui et al. (2019b), Proposition 1. The general definition of $\psi_{\boldsymbol{\alpha}}$ uses conditioning on a σ-lattice, and $\psi_{\boldsymbol{\alpha}}$ is also called a *conditional* 2*-mean* (see Landers and Rogge 1981).

The importance of the function $\psi_{\boldsymbol{\alpha}}$ arises from the fact that we can differentiate this function w.r.t. $\boldsymbol{\alpha}$, in contrast with the least squares estimate $\hat{\psi}_{n,\boldsymbol{\alpha}}$, and that $\psi_{\boldsymbol{\alpha}}$ represents the least squares estimate of ψ_0 in the underlying model for fixed $\boldsymbol{\alpha}$, if we use $\boldsymbol{\alpha}^T\boldsymbol{x}$ as the argument of the monotone link function.

It is also possible to introduce a tuning parameter and use an estimate of $\frac{d}{du}\psi_{\boldsymbol{\alpha}}(u)\big|_{u=\boldsymbol{\alpha}^T\boldsymbol{X}}$. This estimate is defined by

$$\tilde{\psi}'_{n,h,\boldsymbol{\alpha}}(u) = \frac{1}{h}\int K\left(\frac{u-x}{h}\right)\,d\hat{\psi}_{n,\boldsymbol{\alpha}}(x), \tag{6}$$

where K is one of the usual kernels, symmetric around zero and with support $[-1, 1]$, and h is a bandwidth of order $n^{-1/7}$ for sample size n. For fixed $\boldsymbol{\alpha}$, the least squares estimate $\hat{\psi}_{n,\boldsymbol{\alpha}}$ is defined in the same way as above. Note that this estimate is rather different from the derivative of a Nadaraya-Watson estimate which is also used in this context and is in fact the derivative of a ratio of two kernel estimates. If we use the Nadaraya-Watson estimate, we need in principle two tuning parameters, one for the estimation of ψ_0 and another one for the estimation of the derivative ψ_0'.

Using the estimate (6) of the derivative, we now minimize

$$\boldsymbol{\alpha} \mapsto \left\| n^{-1}\sum_{i=1}^{n}\left\{Y_i - \hat{\psi}_{n,\boldsymbol{\alpha}}(\boldsymbol{\alpha}^T\boldsymbol{X}_i)\right\}\boldsymbol{X}_i\,\tilde{\psi}'_{n,h,\boldsymbol{\alpha}}(\boldsymbol{\alpha}^T\boldsymbol{X}_i)\right\|^2 \tag{7}$$

instead of (3), where $\|\cdot\|$ is again the Euclidean norm. The motivation for considering such a minimization problem is very similar to the one given above for the SSE. The only difference now is that the current approach allows us to take the function θ to be equal to the derivative of ψ_0', which is replaced in the empirical version of the population score in (4) by its estimator $\tilde{\psi}'_{n,h,\boldsymbol{\alpha}}$. A variant of this estimator was defined in Balabdaoui et al. (2019b) and called the Efficient Score Estimator (ESE) there, since, if the conditional variance $\mathrm{var}(Y|\boldsymbol{X}=\boldsymbol{x}) = \sigma^2$, where σ^2 is independent of the covariate $\boldsymbol{X}$ (the homoscedastic model), the estimate is efficient. As in the case of the simple score estimator (SSE), the estimate was defined as a crossing of zero estimate in Balabdaoui et al. (2019b) and not as a minimizer of (7). But the definition as a minimizer of (7) produces an estimator that has the same limit distribution.

The qualification "efficient" is somewhat dubious, since the estimator is no longer efficient if we do not have homoscedasticity. We give an example of that situation in Sect. 5, where, in fact, the SSE has a smaller asymptotic variance than the ESE. Nevertheless, to be consistent with our treatment in Balabdaoui et al. (2019b) we will call the estimate, $\hat{\boldsymbol{\alpha}}_n$, minimizing (7), again the ESE.

Dropping the monotonicity constraint, we can also use as our estimator of the link function a cubic spline $\hat{\psi}_{n,\boldsymbol{\alpha}}$, which is defined as the function minimizing

$$\sum_{i=1}^{n} \left\{\psi(\boldsymbol{\alpha}^T \boldsymbol{X}_i) - Y_i\right\}^2 + \mu \int_a^b \psi''(x)^2 \, dx, \tag{8}$$

over the class of functions $\mathcal{S}_2[a, b]$ of differentiable functions ψ with an absolutely continuous first derivative, where

$$a = \min_i \boldsymbol{\alpha}^T \boldsymbol{X}_i, \qquad b = \max_i \boldsymbol{\alpha}^T \boldsymbol{X}_i,$$

see Green and Silverman (1994), pp. 18 and 19, where $\mu > 0$ is the penalty parameter. Using these estimators of the link function, the estimate $\hat{\boldsymbol{\alpha}}_n$ of $\boldsymbol{\alpha}_0$ is then found in Kuchibhotla and Patra (2020) by using a $(d-1)$-dimensional parameterization β and a transformation $S : \boldsymbol{\beta} \mapsto S(\boldsymbol{\beta}) = \boldsymbol{\alpha}$, where $S(\boldsymbol{\beta})$ belongs to the surface of the unit sphere in $\mathbb{R}^d$, and minimizing the criterion

$$\boldsymbol{\beta} \mapsto \sum_{i=1}^{n} \{Y_i - \hat{\psi}_{S(\boldsymbol{\beta}),\mu}(S(\boldsymbol{\beta})^T \boldsymbol{X}_i)\}^2,$$

over $\boldsymbol{\beta}$, where $\hat{\psi}_{S(\boldsymbol{\beta}),\mu}$ minimizes (8) for fixed $\boldsymbol{\alpha} = S(\boldsymbol{\beta})$.

Analogous to our approach above, we can skip the reparameterization and minimize instead

$$\left\| \frac{1}{n} \sum_{i=1}^{n} \left\{\hat{\psi}_{n,\boldsymbol{\alpha},\mu}(\boldsymbol{\alpha}^T \boldsymbol{X}_i) - Y_i\right\} \boldsymbol{X}_i \, \tilde{\psi}'_{n,\boldsymbol{\alpha},\mu}(u)\Big|_{u=\boldsymbol{\alpha}^T \boldsymbol{X}_i} \right\| \tag{9}$$

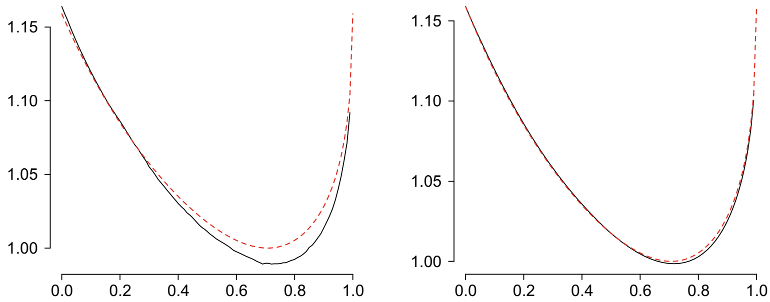
where $\tilde{\psi}_{n,\boldsymbol{\alpha},\mu}$ minimizes (8) for fixed $\boldsymbol{\alpha}$ and $\tilde{\psi}'_{n,\boldsymbol{\alpha},\mu}$ is its derivative. We call this estimator the spline estimator.

We finally give simulation results for these different methods in Sect. 5, where, apart from the comparison with the spline estimator, we make a comparison with other estimators of $\boldsymbol{\alpha}_0$ not using the monotonicity constraint: the Effective Dimension Reduction (EDR) method, proposed in Hristache et al. (2001) and implemented in the `R` package `edr`, the (refined) Mean Average conditional Variance Estimator (MAVE) method, discussed in Xia (2006), and implemented in the `R` package `MAVE`, and Estimation Function Method (EFM), discussed in Cui et al. (2011).

For reasons of space, the proofs of the statements of our paper are given in Balabdaoui and Groeneboom (2020).

2 General Conditions and the Functions $\hat{\psi}_{n,\hat{\alpha}}$ and $\psi_{\hat{\alpha}}$

We give general conditions that we assume to hold in the remainder of the paper here and give graphical comparisons of the functions $\hat{\psi}_{n,\boldsymbol{\alpha}}$ and $\psi_{\boldsymbol{\alpha}}$, where $\psi_{\boldsymbol{\alpha}}$ is defined in Definition 1.



(a) $n = 10^4$ (b) $n = 10^5$

Fig. 1 The loss functions L^{LSE} (red, dashed) and $\widehat{L}_n^{\text{LSE}}$ (solid), where $n = 10^4$ and $n = 10^5$

Example 1 As an illustrative example, we take $d = 2$, $\psi_0(x) = x^3$, $\boldsymbol{\alpha}_0 = (1/\sqrt{2}, 1/\sqrt{2})^T$, $Y_i = \psi_0(\boldsymbol{\alpha}_0^T X_i) + \varepsilon_i$, where the ε_i are i.i.d. standard normal random variables, independent of the $\boldsymbol{X}_i$, which are i.i.d. random vectors, consisting of two independent Uniform(0, 1) random variables. In this case, the conditional expectation function (5) is a rather complicated function of $\boldsymbol{\alpha}$ which we shall not give here but can be computed by a computer package such as Mathematica or Maple. The loss functions:

$$L^{\text{LSE}} : \alpha_1 \mapsto \mathbb{E}\{Y - \psi_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}^T \boldsymbol{X})\}^2 \quad \text{and} \quad \widehat{L}_n^{\text{LSE}} : \alpha_1 \mapsto n^{-1} \sum_{i=1}^{n} \left\{Y_i - \hat{\psi}_{n,\boldsymbol{\alpha}}(\boldsymbol{\alpha}^T \boldsymbol{X}_i)\right\}^2 \tag{10}$$

where the loss function $\widehat{L}_n^{\text{LSE}}$ is for sample sizes $n = 10,000$ and $n = 100,000$, and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)^T$. For $\alpha_1 \in [0, 1]$ and α_2 equal to the positive root $\{1 - \alpha_1^2\}^{1/2}$, we get Fig. 1. The function L^{LSE} has a minimum equal to 1 at $\alpha_1 = 1/\sqrt{2}$, and $\widehat{L}_n^{\text{LSE}}$ has a minimum at a value very close to $1/\sqrt{2}$ (furnishing the profile LSE $\hat{\boldsymbol{\alpha}}_n$), which gives a visual evidence for consistency of the profile LSE.

In order to show the $\sqrt{n}$-consistency and asymptotic normality of the estimators in the next sections, we now introduce some conditions, which correspond to those in Balabdaoui et al. (2019b). We note that we do not need conditions on reparameterization.

(A1) $\boldsymbol{X}$ has a density w.r.t. Lebesgue measure on its support $\mathcal{X}$, which is a convex set $\mathcal{X}$ with a nonempty interior, and satisfies $\mathcal{X} \subset \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\| \le R\}$ for some $R > 0$.
(A2) The function ψ_0 is bounded on the set $\{u \in \mathbb{R} : u = \boldsymbol{\alpha}_0^T \boldsymbol{x},\ \boldsymbol{x} \in \mathcal{X}\}$.
(A3) There exists $\delta > 0$ such that the conditional expectation $\tilde{\psi}_{\boldsymbol{\alpha}}$, defined by (5), is nondecreasing on $I_{\boldsymbol{\alpha}} = \{u \in \mathbb{R} : u = \boldsymbol{\alpha}^T \boldsymbol{x},\ x \in \mathcal{X}\}$ and satisfies $\tilde{\psi}_{\boldsymbol{\alpha}} = \psi_{\boldsymbol{\alpha}}$, so minimizes

$$\left\|\mathbb{E}\left\{Y-\psi(\boldsymbol{\alpha}^T\boldsymbol{X})\right\}\boldsymbol{X}\right\|^2,$$

over nondecreasing functions ψ, if $\|\boldsymbol{\alpha}-\boldsymbol{\alpha}_0\|\le\delta$.

(A4) Let a_0 and b_0 be the (finite) infimum and supremum of the interval $\{\boldsymbol{\alpha}_0^T\boldsymbol{x},\ \boldsymbol{x}\in\mathcal{X}\}$. Then ψ_0 is continuously differentiable on $(a_0-\delta R, a_0+\delta R)$, where R and δ are as in Assumption A1 and A3.

(A5) The density g of $\boldsymbol{X}$ is differentiable and there exist strictly positive constants c_1 to c_4 such that $c_1\le g(\boldsymbol{x})\le c_2$ and $c_3\le\frac{\partial}{\partial x_i}g(\boldsymbol{x})\le c_4$ for $\boldsymbol{x}$ in the interior of $\mathcal{X}$.

(A6) There exists a $c_0>0$ and $M>0$ such that $\mathbb{E}\{|Y|^m|\boldsymbol{X}=\boldsymbol{x}\}\le m!M_0^{m-2}c_0$ for all integers $m\ge 2$ and $\boldsymbol{x}\in\mathcal{X}$ almost surely w.r.t. dG.

These conditions are rather natural, and are discussed in Balabdaoui et al. (2019b). The following lemma shows that, for the asymptotic distribution of $\hat{\boldsymbol{\alpha}}_n$, we can reduce the derivation to the analysis of $\psi_{\hat{\boldsymbol{\alpha}}_n}$. We have the following result (Proposition 4 in Balabdaoui et al. 2019b) on the distance between $\hat{\psi}_{n,\hat{\boldsymbol{\alpha}}}$ and $\psi_{\hat{\boldsymbol{\alpha}}}$.

Lemma 1 *Let conditions (A1)–(A6) be satisfied and let G be the distribution function of* $\boldsymbol{X}$. *Then we have, for* $\boldsymbol{\alpha}$ *in a neighborhood* $\mathcal{B}(\boldsymbol{\alpha}_0,\delta)$ *of* $\boldsymbol{\alpha}_0$

$$\sup_{\boldsymbol{\alpha}\in\mathcal{B}(\boldsymbol{\alpha}_0,\delta)}\int\left\{\hat{\psi}_{n\boldsymbol{\alpha}}(\boldsymbol{\alpha}^T\boldsymbol{x})-\psi_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}^T\boldsymbol{x})\right\}^2 dG(\boldsymbol{x})=O_p\left((\log n)^2 n^{-2/3}\right).$$

3 The Limit Theory for the SSE

In this section, we derive the limit distribution of the SSE introduced above. In our derivation, the function $\psi_{\boldsymbol{\alpha}}$ of Definition 1 plays a crucial role. Below, we will use the following assumptions, additionally to (A1)–(A6).

(A7) There exists a $\delta>0$ such that for all $\boldsymbol{\alpha}\in(\mathcal{B}(\boldsymbol{\alpha}_0,\delta)\cap\mathcal{S}_{d-1})\setminus\{\boldsymbol{\alpha}_0\}$, the random variable

$$\operatorname{cov}\left((\boldsymbol{\alpha}_0-\boldsymbol{\alpha})^T\boldsymbol{X},\,\psi_0(\boldsymbol{\alpha}_0^T\boldsymbol{X})\;\middle|\;\boldsymbol{\alpha}^T\boldsymbol{X}\right)$$

is not equal to 0 almost surely.

(A8) The matrix

$$\mathbb{E}\left[\psi_0'(\boldsymbol{\alpha}_0^T\boldsymbol{X})\operatorname{cov}(\boldsymbol{X}|\boldsymbol{\alpha}_0^T\boldsymbol{X})\right]$$

has rank $d-1$.

We start by comparing (3) with the function

$$\boldsymbol{\alpha}\mapsto\left\|\mathbb{E}\left\{Y-\psi_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}^T\boldsymbol{X})\right\}\boldsymbol{X}\right\|^2. \tag{11}$$