



Nonparametric Multiple-Output Center-Outward Quantile Regression

Eustasio del Barrio, Alberto González Sanz & Marc Hallin

To cite this article: Eustasio del Barrio, Alberto González Sanz & Marc Hallin (20 Sep 2024): Nonparametric Multiple-Output Center-Outward Quantile Regression, Journal of the American Statistical Association, DOI: [10.1080/01621459.2024.2366029](https://doi.org/10.1080/01621459.2024.2366029)

To link to this article: <https://doi.org/10.1080/01621459.2024.2366029>



© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.



[View supplementary material](#)



Published online: 20 Sep 2024.



[Submit your article to this journal](#)



Article views: 690



[View related articles](#)



[View Crossmark data](#)



Citing articles: 2 [View citing articles](#)

Nonparametric Multiple-Output Center-Outward Quantile Regression

Eustasio del Barrio^a, Alberto González Sanz^b, and Marc Hallin^{c,d}

^aIMUVA, Universidad de Valladolid, Valladolid, Spain; ^bDepartment of Statistics, Columbia University, New York, NY; ^cECARES and Département de Mathématique, Université libre de Bruxelles, Bruxelles, Belgium; ^dInstitute of Information Theory and Automation, Czech Academy of Sciences, Prague, Czech Republic

ABSTRACT

Building on recent measure-transportation-based concepts of multivariate quantiles, we are considering the problem of nonparametric multiple-output quantile regression. Our approach defines nested conditional center-outward quantile regression contours and regions with given conditional probability content, the graphs of which constitute nested center-outward *quantile regression tubes* with given unconditional probability content; these (conditional and unconditional) probability contents do not depend on the underlying distribution—an essential property of quantile concepts. Empirical counterparts of these concepts are constructed, yielding interpretable empirical contours, regions, and tubes which are shown to consistently reconstruct (in the Pompeiu-Hausdorff topology) their population versions. Our method is entirely nonparametric and performs well in simulations—with possible heteroscedasticity and nonlinear trends. Its potential as a data-analytic tool is illustrated on some real datasets. Supplementary materials for this article are available online, including a standardized description of the materials available for reproducing the work.

ARTICLE HISTORY

Received December 2022
Accepted May 2024

KEYWORDS

Center-outward quantiles;
Multiple-output regression;
Optimal transport

1. Introduction



1.1. Quantile Regression, Single- and Multiple-Output

Forty-five years after its introduction by Koenker and Bassett (1978), quantile regression—arguably the most powerful tool in the statistical study of the dependence of a variable of interest Y on covariates $\mathbf{X} = (X_1, \dots, X_m)$ —has become part of statistical daily practice, with countless applications in all domains of scientific research, from economics and social sciences to astronomy, biostatistics, and medicine. Unlike classical regression which, somewhat narrowly, is focused on the conditional means $E[Y|\mathbf{X} = \mathbf{x}]$, quantile regression indeed is dealing with the complete conditional distributions $P_{Y|\mathbf{X}=\mathbf{x}}$ of Y conditional on $\mathbf{X} = \mathbf{x}$. Building on that pioneering contribution, a number of quantile regression methods, parametric, semiparametric, and nonparametric, have been developed for an extremely broad range of statistical topics, including time series, survival analysis, instrumental variables, measurement errors, and functional data—to quote only a few. Sometimes, a simple parameterized regression model allows for a parametric approach, yielding, for instance, linear quantile regression. In most situations, however, parametric models are too rigid and a more agnostic nonparametric approach is in order. We refer to Koenker (2005) for an introductory text, to Koenker et al. (2018) for a comprehensive survey.

In single-output models (univariate variable of interest Y), this nonparametric approach is well understood and well studied, and the history of nonparametric estimation of conditional quantile functions goes back, at least, to the seminal paper

by Stone (1977). The results are much scarcer, however, in the ubiquitous multiple-output case (d -dimensional variable of interest \mathbf{Y} , with $d > 1$), and sometimes less convincing—the simple reason for this being the problem of defining a fully satisfactory concept of multivariate quantiles. A major difficulty with quantiles in dimension $d > 1$, indeed, is the fact that \mathbb{R}^d , contrary to \mathbb{R} , is not canonically ordered. A number of attempts have been made to overcome that issue, the most remarkable of which are the theory of statistical depth and the concept of geometric or spatial quantiles.

The theory of statistical depth has generated an abundant literature which we cannot summarize here—we refer to (Serfling and Zuo 2000) or (Serfling 2002, 2019) for general expositions and surveys. Several depth concepts coexist. The most popular of them is Tukey's halfspace depth (Tukey 1975), but all depth concepts (including the most recent ones: see, e.g., Koen and Paindaveine (2022)) are sharing the same basic properties. Tukey's halfspace depth characterizes, for each distribution P over \mathbb{R}^d (for simplicity, assume P to be Lebesgue-absolutely continuous) *depth regions* $\mathbb{D}_P(\delta)$ (resp., *depth contours* $\mathcal{D}_P(\delta)$) as collections of points with depth (relative to P) larger than or equal to δ (resp., equal to δ), $\delta \in (0, 1/2]$. Depth regions, moreover, are convex (even for distributions with highly nonconvex shapes), closed, connected, and nested as δ increases. Depth regions were proposed as notions of quantile regions—an interpretation that is supported, for example, by the L_1 nature of Tukey depth (Hallin, Paindaveine, and Šiman 2010). Among the merits of this interpretation of depth regions as quantile

CONTACT Marc Hallin  mhallin@ulb.ac.be  ECARES and Département de Mathématique, Université libre de Bruxelles, Avenue Franklin Roosevelt 50, 1050 Bruxelles, Belgium.

 Supplementary materials for this article are available online. Please go to www.tandfonline.com/r/JASA.

© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

regions is the fact that it has imposed the idea that quantiles, in dimension $d \geq 2$, should rely on some center-outward ordering (with central region the deepest points) rather than a southwest-northeast extension of the classical univariate “left-to-right” linear ordering of the real line.

Unfortunately, depth regions do not characterize the underlying distribution (see, e.g., Nagy 2021) and fail to satisfy the quintessential property of quantile regions: the P-probability content $P[\mathbb{D}_P(\delta)]$ of the quantile region $\mathbb{D}_P(\delta)$ indeed very much depends on P.¹ None of the depth concepts in the literature is enjoying the essential property of a P-probability content $P[\mathbb{D}_P(\delta)]$ independent of P, though, and depth regions, therefore, cannot be considered as *bona fide* quantile regions. A remedy that is often proposed to that problem consists in reindexing depth regions $\mathbb{D}_P(\delta)$ and contours $\mathcal{D}_P(\delta)$ with their P-probability contents $P(\mathbb{D}_P(\delta))$. Such reindexation, however, can be performed with *any* collection of nested regions, turning into quantile regions an arbitrary collection of sets bearing no quantile-type relation,² or even no relation at all, to P.

A more attractive concept of multivariate quantiles, which does (Koltchinski 1997; Konen 2023) characterize the underlying distribution P, is provided by the L_1 definition of *geometric* or *spatial* quantiles introduced by Chaudhuri (1996) (see Koltchinski (1996) for a generalization and, for extensions to functional variables, Chakraborty and Chaudhuri 2014; Chowdhury and Chaudhuri 2019). The same reindexing procedure as for depth-based quantiles is required, however, if the probability of geometric quantile regions are to be independent of P and, even so, the inverse of the reindexed geometric quantile function of P (in principle, a *distribution function* for P), unlike the measure-transportation-based one to be considered below, fails to be distribution-free when computed at a random variable with distribution P—while such distribution-freeness is a typical feature of distribution functions, which by nature are probability integral transformations; Girard and Stupfler (2017), moreover, stress a few unpleasant properties of extreme geometric quantiles (see also Hallin and Konen 2024). Projection methods also have been used to define multivariate quantile regions based on univariate directional ones (Kong and Mizera 2012; Paindaveine and Šiman 2012), mostly leading to depth-based quantile contours yielding the same weaknesses as described above.

In the presence of covariates, each multivariate quantile concept quite naturally yields a conditional version which in turn suggests a multiple-output quantile regression model. A review of various existing multiple-output quantile regression models (linear and nonparametric, depth-based, and others) can be found in Hallin and Šiman (2018). Nonparametric quantile regression models based on Tukey depth are developed in Hallin et al. (2015) but suffer from the same lack of control over

the probability contents of the resulting (conditional) quantile regions as the depth-based quantile regions themselves. So do the directional concepts of M-quantiles (Breckling and Chambers 1988; Paindaveine and Šiman 2011; Merlo et al. 2022) and the projection-based one in Wei (2008). The reindexed geometric quantile regression approach (Chaudhuri 1996, sec. 2.2; Chakraborty 2003; Chowdhury and Chaudhuri 2019 for functional versions) offers such a control but suffer from the same major drawback (Girard and Stupfler 2017) as their unconditional version.

1.2. Measure-Transportation-based Quantile Regression

Recently, based on measure transportation ideas, new concepts of depth and quantiles in dimension $d > 1$ have been introduced under the names of *Monge-Kantorovich depth* and *center-outward quantile function* by Chernozhukov et al. (2017) and Hallin et al. (2021). The idea consists in defining the quantile function of P as the a.s. unique gradient of convex function *pushing* some uniform reference distribution *forward*³ to P. In \mathbb{R}^d , the reference distribution usually is chosen as the uniform $U_{[0,1]^d}$ over the unit cube $[0, 1]^d$ or the spherical uniform U_d over the unit ball \mathbb{S}^d . The advantage of the latter is that, due to spherical symmetry, the collection $\{\tau\bar{\mathbb{S}}^d, \tau \in [0, 1]\}$ of closed nested balls constitutes for U_d a natural collection of central quantile regions (note that $U_d(\tau\bar{\mathbb{S}}^d) = \tau$). This is the choice we tacitly adopt throughout this article, calling center-outward quantile function the a.s. unique gradient of convex function \mathbf{Q}_\pm such that $P = \mathbf{Q}_\pm \# U_d$: see Hallin et al. (2021) for details and Hallin (2022) for a survey of measure-transportation-based center-outward quantiles, the dual concepts of center-outward distribution functions, and their empirical counterparts—multivariate ranks and signs—along with a list of applications in inference (Shi et al. 2021; Ghosal and Sen 2022; Hallin, La Vecchia, and Liu 2022; Deb and Sen 2023; Hallin and Liu 2023; Hallin, Hlubinka, and Hudcová 2023; Hallin, La Vecchia, and Liu 2023; Hallin and Mordant 2023; Shi et al. 2024, etc.).

Center-outward quantile functions characterize nested, closed, and connected quantile regions $\mathbb{C}_P(\tau) := \mathbf{Q}_\pm(\tau\bar{\mathbb{S}}^d)$ and continuous contours $\mathcal{C}_P(\tau) := \mathbf{Q}_\pm(\tau\bar{\mathbb{S}}^d)$ indexed by $\tau \in (0, 1)$ such that, for any absolutely continuous P, $P[\mathbb{C}_P(\tau)] = \tau$ irrespective of P. These measure-transportation-based quantiles thus do satisfy the essential property that the P-probability contents of the resulting quantile regions do not depend on P. Moreover, these regions are not necessarily convex and, as shown in Figure 1, they are able to capture the “shape” of the underlying distribution. Finally, the inverse \mathbf{F}_\pm of \mathbf{Q}_\pm is a probability integral transformation, that is, pushes P forward to a uniform distribution.

The notion of quantile region is less clear for $U_{[0,1]^d}$; one could consider the centered cube $[\frac{1}{2} \pm \frac{1}{2}\tau^{-d}]^d$ (with $U_{[0,1]^d}$ -probability content τ) as a central quantile region of order τ for $U_{[0,1]^d}$, and transport it to \mathbb{R}^d via the a.s. unique gradient of

¹This is not a minor weakness: the univariate median $Y_{1/2}^P$ of an absolutely continuous distribution P, for instance, is characterized by the fact that $P((-\infty, Y_{1/2}^P]) = 1/2$ irrespective of P—who would call *median* a quantity Y_{med}^P such that $P_1((-\infty, Y_{\text{med}}^P]) = 0.4$ for some P_1 , while $P_2((-\infty, Y_{\text{med}}^P]) = 0.6$ for some P_2 ?

²Reindexation by probability content is used, for example, by Camehl, Fok, and Gruber (2022) for density level contours, which are not necessarily connected hence do not enjoy the essential monotonicity features of quantiles.

³Recall that a mapping T is *pushing* P_1 forward to P_2 (notation: $T\#P_1 = P_2$) if $T(\mathbf{X}) \sim P_2$ when $\mathbf{X} \sim P_1$. For Lebesgue-absolutely continuous distributions on \mathbb{R}^d , a famous theorem by McCann (1995) establishes the existence and a.s. uniqueness of a gradient of convex function T such that $T\#P_1 = P_2$.

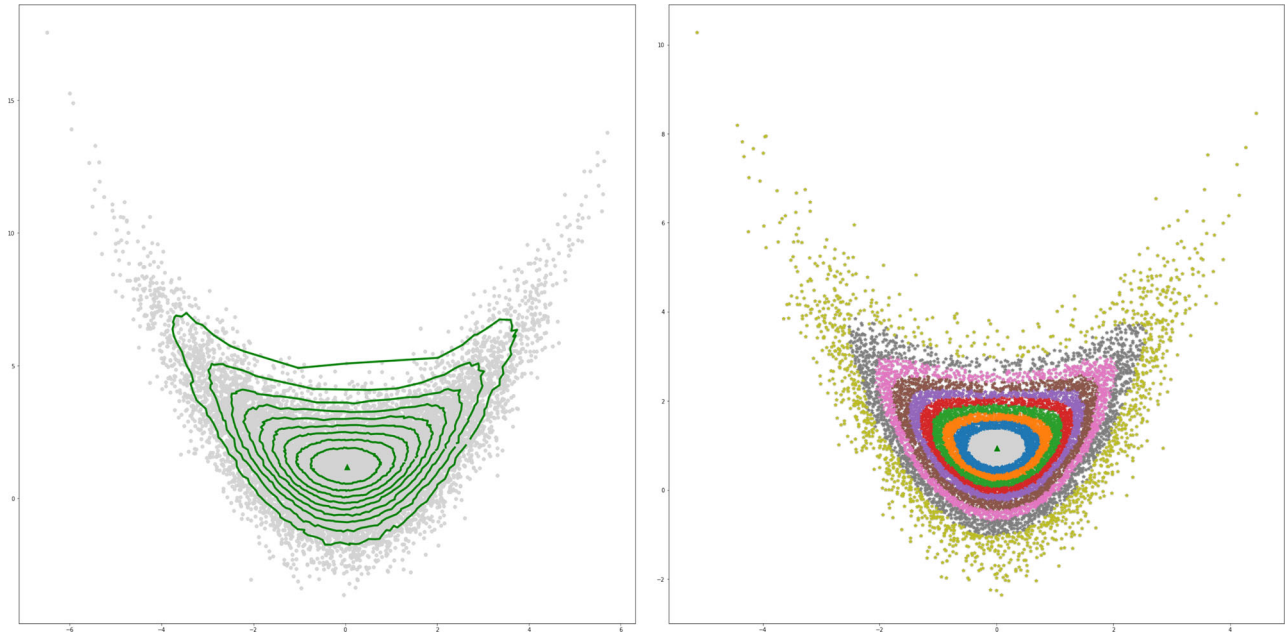


Figure 1. Quantile contours $C_{\pm}(\tau)$ (left panel) and regions $\mathbb{C}_{\pm}(\tau)$ (right panel) for the Gaussian mixture of Section 4.1.2 and quantile orders $\tau = 0.1, 0.2, \dots, 0.9$.

convex function \mathbf{Q}_{\square} pushing $U_{[0,1]^d}$ forward to P , thus, yielding quantile regions of the form $\mathbf{Q}_{\square}([\frac{1}{2} \pm \frac{1}{2}\tau^{-d}]^d)$. Another choice is the image by \mathbf{Q}_{\square} of region of the unit cube lying under the hyperboloid $\prod_{i=1}^d u_i = \tau_d$ where τ_d is such that the region has volume τ . None of these choices is preserving the symmetry features of P : in particular, the resulting quantile regions of a spherical distribution fail to be spherically symmetric. Based on the conditional version of \mathbf{Q}_{\square} , Carlier, Chernozhukov, and Galichon (2016) are proposing a quantile regression model focused on linear regression which does not yield any obvious (conditional) quantile regions and contours, though. A functional extension of the same approach is developed in Agarwal et al. (2022).

Motivated by the long list of successful applications of center-outward quantiles, ranks, and signs, we are proposing in this article a novel and meaningful solution, based on the concept of center-outward quantiles, to the problem of nonparametric multiple-output quantile regression. Namely, for a pair of multidimensional random variables (\mathbf{X}, \mathbf{Y}) with values in $\mathbb{R}^m \times \mathbb{R}^d$ (\mathbf{Y} the variable of interest, \mathbf{X} the vector of covariates) and joint distribution⁴ \mathbb{P} , we define (Section 2.2) the *center-outward quantile function of the conditional distribution of \mathbf{Y} conditional on $\mathbf{X} = \mathbf{x}$* as

$$\mathbf{Q}_{\pm}(\cdot | \mathbf{x}) : \mathbf{u} \in \mathbb{S}_d \mapsto \mathbf{Q}_{\pm}(\mathbf{u} | \mathbf{x}) \in \mathbb{R}^d \quad (1.1)$$

(\mathbb{S}_d the open unit ball in \mathbb{R}^d), with the essential property that, letting

$$\mathbb{C}_{\pm}(\tau | \mathbf{x}) := \mathbf{Q}_{\pm}(\tau \bar{\mathbb{S}}_d | \mathbf{x}) \quad \tau \in (0, 1), \quad \mathbf{x} \in \mathbb{R}^m, \quad (1.2)$$

we have, irrespective of \mathbb{P} ,

$$\mathbb{P}[\mathbf{Y} \in \mathbb{C}_{\pm}(\tau | \mathbf{x}) | \mathbf{X} = \mathbf{x}] = \tau \quad \text{for all } \mathbf{x} \in \mathbb{R}^m, \text{ and } \tau \in (0, 1), \quad (1.3)$$

justifying the interpretation of $\mathbf{x} \mapsto \mathbb{C}_{\pm}(\tau | \mathbf{x})$ as the value at \mathbf{x} of a *regression quantile region of order τ* of \mathbf{Y} conditional on $\mathbf{X} = \mathbf{x}$. For $\tau = 0$,

$$\mathbb{C}_{\pm}(0 | \mathbf{x}) := \bigcap_{\tau \in (0,1)} \mathbb{C}_{\pm}(\tau | \mathbf{x}) \quad (1.4)$$

yields the value at $\mathbf{X} = \mathbf{x}$ of the *regression median* $\mathbf{x} \mapsto \mathbb{C}_{\pm}(0 | \mathbf{x})$ of \mathbf{Y} conditional on $\mathbf{X} = \mathbf{x}$. The same conditional quantile map characterizes nested (no “quantile crossing” phenomenon) *regression quantile tubes of order τ* (in \mathbb{R}^{m+d})

$$\mathbb{T}_{\pm}(\tau) := \left\{ (\mathbf{x}, \mathbf{Q}_{\pm}(\tau \bar{\mathbb{S}}_d | \mathbf{x})) \mid \mathbf{x} \in \mathbb{R}^m \right\}, \quad \tau \in (0, 1) \quad (1.5)$$

which are such that

$$\mathbb{P}[(\mathbf{X}, \mathbf{Y}) \in \mathbb{T}_{\pm}(\tau)] = \tau \quad \text{irrespective of } \mathbb{P}, \tau \in (0, 1). \quad (1.6)$$

For $\tau = 0$, define

$$\mathbb{T}_{\pm}(0) := \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \in \mathbb{R}^m, \mathbf{y} \in \mathbb{C}_{\pm}(0 | \mathbf{x})\} = \bigcap_{\tau \in (0,1)} \mathbb{T}_{\pm}(\tau)$$

(the *graph* of $\mathbf{x} \mapsto \mathbb{C}_{\pm}(\tau | \mathbf{x})$); with a slight abuse of language, also call $\mathbb{T}_{\pm}(0)$ the *regression median* of \mathbf{Y} with respect to \mathbf{X} .

Except for the reindexed ones,⁵ none of the earlier attempts to define multiple-output regression quantiles is characterizing quantile regions that satisfy requirements (1.3) and (1.6). As the first attempt to break with directional and depth-based approaches to multiple-output quantile regression by means of innovative measure transportation ideas, Carlier, Chernozhukov, and Galichon (2016) deserves special attention. While providing (under linear regression assumptions) a consistent reconstruction of the distributions of \mathbf{Y} conditional on $\mathbf{X} = \mathbf{x}$, however, they do not propose any definition of

⁴For simplicity, we tacitly assume all distributions to be Lebesgue-absolutely continuous.

⁵A typical case is the approach proposed by Camehl, Fok, and Gruber (2022) which, however, qualifies as a density level regression method rather than a quantile regression one.

quantile regions with given probability content τ similar to the quantile regions $\mathbb{C}_{\pm}(\tau|\mathbf{x}) \subset \mathbb{R}^d$ or the quantile regression tubes $\mathbb{T}_{\pm}(\tau) \subset \mathbb{R}^{m+d}$. The validity of the estimation method they are proposing, and their derivation of asymptotic results, moreover, is limited to *linear* quantile regression and focused on a concept of quantile function defined over the unit cube $[0, 1]^d$ rather than the unit ball \mathbb{S}_d .

1.3. Motivating Examples

Figure 2 provides, for $m = 1$ and $d = 2$, a visualization of the regression median and quantile tubes of orders $\tau = 0.2, 0.4$, and 0.8 , along with some of the corresponding *conditional regression quantile contours* $\mathbb{C}_{\pm}(\tau|\mathbf{x}) := \{(\mathbf{x}, \mathbf{Q}_{\pm}(\tau \mathcal{S}_{d-1}|\mathbf{x}))\}$ (\mathcal{S}_{d-1} the unit sphere in \mathbb{R}^d), for two models. In the first example (left panel), (X, \mathbf{Y}) are of the form

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} \sin(\frac{2\pi}{3}X) + 0.575\sigma(X)e_1 \\ \cos(\frac{2\pi}{3}X) + X^2 + 2.65X^4 + \frac{\sigma(X)e_2}{2.3} + \frac{1}{4}e_1^3 \end{pmatrix} \quad (1.7)$$

where $X \sim U_{[-1,1]}$, $\sigma(X) = (1 + \frac{3}{2}\sin(\pi X/2))^2$, and $\mathbf{e} = (e_1, e_2)^T \sim \mathcal{N}(\mathbf{0}, \mathbf{Id})$, with X and \mathbf{e} mutually independent. In the second example (right panel),

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X \\ X^2 \end{pmatrix} + \left(1 + \frac{3}{2}\sin\left(\frac{\pi}{2}X\right)^2\right)\mathbf{R}(X)\mathbf{e}, \quad (1.8)$$

where $X \sim U_{[-2,2]}$ and \mathbf{e} are mutually independent, $10\mathbf{e}$ is a mixture (equal mixing probabilities 0.25) of $\mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{Id}\right)$, $\mathcal{N}\left(\begin{pmatrix} 8.66 \\ -5.00 \end{pmatrix}, \mathbf{Id}\right)$, $\mathcal{N}\left(\begin{pmatrix} -8.66 \\ -5.00 \end{pmatrix}, \mathbf{Id}\right)$, and $\mathcal{N}\left(\begin{pmatrix} 0 \\ 10 \end{pmatrix}, \mathbf{Id}\right)$, and $\mathbf{R}(x)$ is the rotation matrix of angle $\frac{\pi}{2}x$.

Actually, since explicit values cannot be obtained for these population concepts, very large samples of $n = 100,000$ observations were generated from (1.7) and (1.8), respectively, and the consistent estimation procedure described in Section 3 was performed to obtain the pictures in Figure 2. Note the non-convexity of the conditional contours, the non-linearity of the

regression median, and the marked heteroscedasticity of the regression. Also note that our method is numerically implementable for very large n .

1.4. Outline of the Article

The article is organized as follows. Section 2 is dealing with the population concept of *conditional center-outward quantile map* and the resulting *center-outward regression quantile contours and regions*, Section 3 with their estimation. In Section 3.1, we show how to construct empirical quantile contours and regions, the consistency of which is established in Section 3.2. Numerical results are provided in Section 4. Monte Carlo experiments in Section 4.1 show the ability of our method to handle very large sample sizes and to pick heteroscedasticity, nonlinear trends, and the shape of conditional distributions; comparisons also are made with the results of Hallin et al. (2015) and Carlier, Chernozhukov, and Galichon (2016). Real datasets are analyzed in Section 4.3, illustrating the power of our method as a data-analytic tool. Section 5 concludes with some recent references on the numerical aspects of optimal transports and perspectives for future developments. An online supplement provides mathematical proofs, further discussion of computational issues, and further numerical results.

2. Nonparametric Center-Outward Quantile Regression

2.1. Notation

For convenience, we are listing here the main notation to be used throughout the article. Unless otherwise stated, we denote by $(\Omega, \mathcal{A}, \mathbb{P})$ the triple defining the underlying probability space. Let ℓ_d denote the d -dimensional Lebesgue measure, \mathcal{B}_d the Borel σ -field, and $\mathcal{P}(\mathbb{R}^d)$ the space of Borel probability measures on \mathbb{R}^d . The support of $\mathbb{P} \in \mathcal{P}(\mathbb{R}^d)$ is denoted as $\text{supp}(\mathbb{P})$. Throughout, (X, \mathbf{Y}) denotes an \mathbb{R}^{m+d} -valued random vector

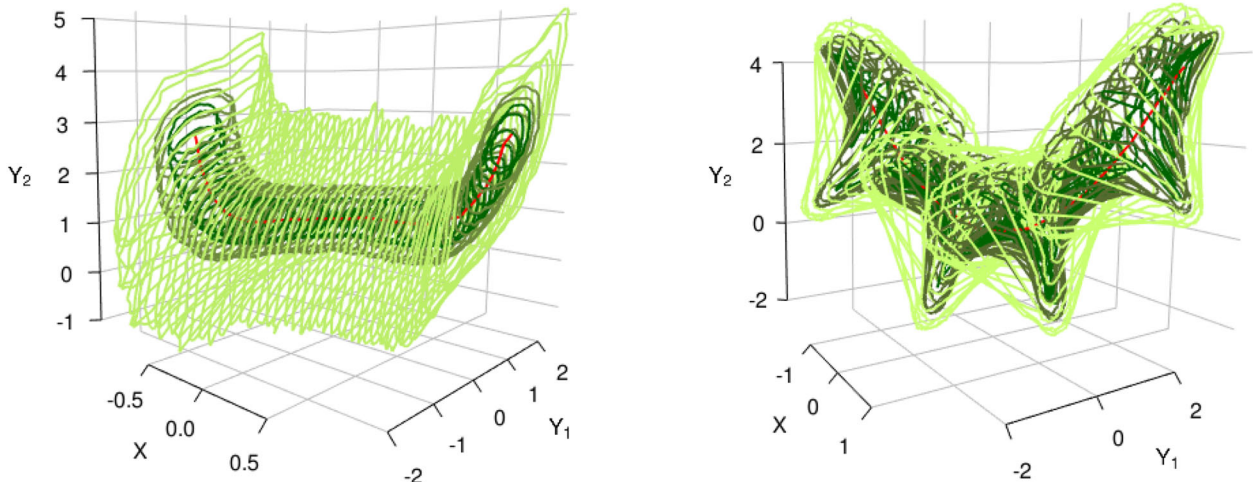


Figure 2. Two examples ((1.7), (1.8)) of multiple-output quantile regression tubes (bivariate variable of interest $\mathbf{Y} = (Y_1, Y_2)$; univariate regressor X), showing the conditional center-outward medians (red) and the conditional quantile contours $(\mathbf{x}, \mathbb{C}_{\pm}(\tau|\mathbf{x}))$ of order $\tau = 0.2$ (black), $\tau = 0.4$ (dark green), and $\tau = 0.8$ (light green). Numerical approximations based on $n = 100,000$ points.

with distribution $\mathbb{P} = P_{\mathbf{X},\mathbf{Y}} \in \mathcal{P}(\mathbb{R}^{m+d})$, m -dimensional \mathbf{X} -marginal $P_{\mathbf{X}}$ and d -dimensional \mathbf{Y} -marginal $P_{\mathbf{Y}}$. The distribution of \mathbf{Y} conditional on $\mathbf{X} = \mathbf{x}$ is denoted as $P_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}$.⁶ The open and closed unit ball and the unit hypersphere in \mathbb{R}^d are denoted by $\mathbb{S}_d, \bar{\mathbb{S}}_d$, and \mathcal{S}_{d-1} , respectively. We denote by U_d the spherical uniform over \mathbb{S}_d —that is, the product of a uniform distribution V_d over the unit hypersphere \mathcal{S}_{d-1} (for the directions) and a uniform distribution over $[0, 1]$ (for the distance to the center).

2.2. Conditional Center-Outward Quantiles, Regions, and Contours

Let us provide precise definitions for the concepts we briefly presented in the Introduction and properly introduce center-outward quantiles, regions, and contours.

For any $P \in \mathcal{P}(\mathbb{R}^d)$, denote by $\mathbf{Q}_{\pm} = \nabla\varphi$ and call *center-outward quantile map* the (Lebesgue-a.e.) unique gradient of a convex function $\varphi : \mathbb{S}_d \rightarrow \mathbb{R}$ such that $\mathbf{Q}_{\pm}(U) \sim P$ for any $U \sim U_d$ —in the measure transportation convenient terminology, \mathbf{Q}_{\pm} is *pushing* P forward to U_d , which we denote as $\mathbf{Q}_{\pm} \# P = U_d$. This only defines \mathbf{Q}_{\pm} at φ 's points of differentiability (recall that convex functions are differentiable at almost every point in the interior of their domain: see Theorems 26.1 and 25.5 in Rockafellar 1970). At φ 's points of non-differentiability \mathbf{u} , let us define $\mathbf{Q}_{\pm}(\mathbf{u})$ as the subdifferential $\partial\varphi(\mathbf{u})$ of φ , namely,

$$\mathbf{Q}_{\pm}(\mathbf{u}) := \partial\varphi(\mathbf{u}) := \left\{ \mathbf{y} \in \mathbb{R}^d \mid \text{for all } \mathbf{z} \in \mathbb{R}^d, \right. \\ \left. \varphi(\mathbf{z}) - \varphi(\mathbf{u}) \geq \langle \mathbf{y}, \mathbf{z} - \mathbf{u} \rangle \right\}, \quad \mathbf{u} \in \mathbb{S}_d;$$

then, \mathbf{Q}_{\pm} is an everywhere-defined set-valued function. Slightly abusing notation, we also write \mathbf{Q}_{\pm} and $\partial\varphi$ for the set of all points $(\mathbf{u}, \mathbf{y}) \in \mathbb{R}^{m+d}$ such that $\mathbf{y} \in \partial\varphi(\mathbf{u})$. We then can introduce the concepts of conditional center-outward quantiles, contours, and regions.

Definition 2.1. Call *conditional center-outward quantile map of \mathbf{Y} given $\mathbf{X} = \mathbf{x}$* the center-outward quantile map

$$\mathbf{u} \mapsto \mathbf{Q}_{\pm}(\mathbf{u}|\mathbf{X} = \mathbf{x}), \mathbf{u} \in \mathbb{S}_d$$

of $P_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}$, $\mathbf{x} \in \mathbb{R}^m$. The corresponding *conditional center-outward quantile regions and contours of order $\tau \in (0, 1)$* are the sets $\mathbb{C}_{\pm}(\tau|\mathbf{x}) := \mathbf{Q}_{\pm}(\tau\bar{\mathbb{S}}_d|\mathbf{X} = \mathbf{x})$ and $\mathbb{C}_{\pm}(\tau|\mathbf{x}) := \mathbf{Q}_{\pm}(\tau\mathcal{S}_{d-1}|\mathbf{X} = \mathbf{x})$, respectively. The conditional center-outward quantile maps also characterize (see Definitions (1.4), (1.5), and (1.6)) *conditional medians* $\mathbb{C}_{\pm}(0|\mathbf{x})$ and *regression quantile tubes* $\mathbb{T}_{\pm}(\tau)$.

When no confusion is possible, we also write $\mathbf{Q}_{\pm}(\mathbf{u}|\mathbf{x})$ for $\mathbf{Q}_{\pm}(\mathbf{u}|\mathbf{X} = \mathbf{x})$. The terminology center-outward *regression quantile region, contour, and median* is used for the mappings $\mathbf{x} \mapsto \mathbb{C}_{\pm}(\tau|\mathbf{x})$, $\mathbf{x} \mapsto \mathbb{C}_{\pm}(\tau|\mathbf{x})$, and $\mathbf{x} \mapsto \mathbb{C}_{\pm}(0|\mathbf{x})$, respectively, $\mathbf{x} \in \mathbb{R}^m$.

Recall, however, that, in the absence of any assumptions on the conditional probabilities $P_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}$, the mappings $\mathbf{u} \mapsto \mathbf{Q}_{\pm}(\mathbf{u}|\mathbf{X} = \mathbf{x})$ typically are set-valued, see Rockafellar and Wets

(1998). Whenever continuous, single-valued functions (typically, on the punctured unit ball $\mathbb{S}_d \setminus \{0\}$) are needed, we will make the following assumption.

Assumption (R). The conditional distribution $P_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}$ admits, $P_{\mathbf{X}}$ -a.e. $\mathbf{x} \in \mathbb{R}^m$, a convexly supported density $p_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}$ with respect to the Lebesgue measure. Moreover, that density $p_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}$ is such that, for every $R > 0$ and \mathbf{x} , there exist constants $0 < \lambda_R^{\mathbf{x}} \leq \Lambda_R^{\mathbf{x}} < \infty$ such that $\lambda_R^{\mathbf{x}} \leq p_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}(\mathbf{y}) \leq \Lambda_R^{\mathbf{x}}$ for all $\mathbf{y} \in \text{supp}(P_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}) \cap R\mathbb{S}_d$.

In the classical single-output case ($d = 1$), consistent estimation of conditional quantiles similarly requires the continuity of the conditional quantile maps (see Stone 1977). In dimension $d > 1$, the continuity of center-outward quantile maps follows from assumptions similar to Assumption (R)—see Figalli (2018) and del Barrio, González-Sanz, and Hallin (2020).

3. Empirical Center-Outward Quantile Regression

We now proceed with the construction of empirical versions of the conditional center-outward quantile concepts defined in Section 2.2 and their consistency properties.

3.1. Empirical Conditional Center-Outward Quantiles

Let $(\mathbf{X}, \mathbf{Y})^{(n)} := ((\mathbf{X}_1, \mathbf{Y}_1), \dots, (\mathbf{X}_n, \mathbf{Y}_n))$ be a sample of n iid copies of $(\mathbf{X}, \mathbf{Y}) \sim \mathbb{P}P_{\mathbf{X},\mathbf{Y}}$. In this section, we develop an estimator of the conditional center-outward quantile maps $\mathbf{u} \mapsto \mathbf{Q}_{\pm}(\mathbf{u}|\mathbf{X} = \mathbf{x})$, $\mathbf{x} \in \mathbb{R}^m$. Our estimator is obtained in two steps: in Step 1, we construct an empirical distribution of \mathbf{Y} conditional on $\mathbf{X} = \mathbf{x}$ and, in Step 2, we compute the corresponding empirical center-outward quantile map.

Step 1. For each value of $\mathbf{x} \in \mathbb{R}^m$, our estimation of the distribution of \mathbf{Y} conditional on $\mathbf{X} = \mathbf{x}$ involves a sequence of *weight functions* $w^{(n)} : \mathbb{R}^{m(n+1)} \rightarrow \mathbb{R}^n$ of the form

$$(\mathbf{x}, \mathbf{X}^{(n)}) \mapsto w^{(n)}(\mathbf{x}, \mathbf{X}^{(n)}) := (w_1(\mathbf{x}; \mathbf{X}^{(n)}), \dots, w_n(\mathbf{x}; \mathbf{X}^{(n)}))$$

where $w_j^{(n)} : \mathbb{R}^{m(n+1)} \rightarrow \mathbb{R}$ is measurable with respect to \mathbf{x} and the sample $\mathbf{X}^{(n)} := (\mathbf{X}_1, \dots, \mathbf{X}_n)$ and satisfies

$$w_j^{(n)}(\mathbf{x}; \mathbf{X}^{(n)}) \geq 0 \quad \text{and}$$

$$\sum_{j=1}^n w_j^{(n)}(\mathbf{x}; \mathbf{X}^{(n)}) = 1 \quad \text{a.s. for all } n, \quad j = 1, \dots, n. \quad (3.1)$$

We refer to a function $w^{(n)}$ satisfying (3.1) as a *probability weight function* and define the *empirical conditional distribution of \mathbf{Y} given $\mathbf{X} = \mathbf{x}$* as $P_{w(\mathbf{x})}^{(n)} := \sum_{j=1}^n w_j^{(n)}(\mathbf{x}; \mathbf{X}^{(n)})\delta_{\mathbf{Y}_j}$ where $\delta_{\mathbf{Y}_j}$ is the Dirac function computed at \mathbf{Y}_j . Following Stone (1977), we say that the sequence $w^{(n)}$ is a *consistent weight function* if, whenever $(\mathbf{X}, \mathbf{Y}), (\mathbf{X}_1, \mathbf{Y}_1), \dots, (\mathbf{X}_n, \mathbf{Y}_n)$ are iid, where \mathbf{Y} is real-valued and such that $E|\mathbf{Y}|^r < \infty$ for some $r > 1$,

$$E \left| \sum_{j=1}^n w_j^{(n)}(\mathbf{X}; \mathbf{X}^{(n)}) \mathbf{Y}_j - E(\mathbf{Y}|\mathbf{X}) \right|^r \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (3.2)$$

Step 2. To estimate the conditional quantiles, consider a *regular grid* $\mathfrak{G}^{(N)}$ of \mathbb{S}_d consisting of N gridpoints denoted as

⁶The existence of the regular conditional probability is a direct consequence of the so-called *disintegration theorem* (see, e.g., Theorem 2.5.1 in Lehmann and Romano 2005).

$\mathfrak{G}_1^{(N)}, \dots, \mathfrak{G}_N^{(N)}$. The number N here is arbitrarily chosen as factorizing into a product of integers of the form $N = N_R N_S + N_0$ with $N_0 = 0$ or 1 . That regular grid is created as the intersection between

- the rays generated by an N_S -tuple $\mathbf{u}_1, \dots, \mathbf{u}_{N_S} \in \mathcal{S}_{d-1}$ of unit vectors such that $N_S^{-1} \sum_{j=1}^{N_S} \delta_{\mathbf{u}_j}$ converges weakly to the uniform over \mathcal{S}_{d-1} as $N_S \rightarrow \infty$, and
- the N_R hyperspheres with center $\mathbf{0}$ and radii $j/(N_R + 1)$, $j = 1, \dots, N_R$,

along with the origin if $N_0 = 1$. Based on this grid, we define the sequence of discrete uniform measures

$$U_d^{(N)} := \frac{1}{N} \sum_{j=1}^N \delta_{\mathfrak{G}_j^{(N)}} \in \mathcal{P}(\mathbb{R}^d), N \in \mathbb{N}$$

over $\mathfrak{G}^{(N)}$ and require that both $N_R \rightarrow \infty$ and $N_S \rightarrow +\infty$ as $N \rightarrow \infty$. By construction, $U_d^{(N)}$ converges weakly to U_d as $N \rightarrow \infty$. Note that imposing $N_0 = 0$ or 1 is not a problem, since N , unlike n , is chosen by the practitioner; it yields the fundamental advantage that all points of $\mathfrak{G}^{(N)}$ have multiplicity one so that Corollary 3.1 in Hallin et al. (2021), to be used below, applies.

Our estimation of the conditional center-outward quantile maps relies on the optimal transport pushing $U_d^{(N)}$ forward to $P_{w(\mathbf{x})}^{(n)}$ —more precisely, adopting (since typically $N \neq n$) the Kantorovich formulation of the optimal transport problem, on the solution of the linear program (solvable using efficient numerical methods such as the auction or Hungarian algorithms—see Peyré and Cuturi (2019) and references therein)⁷

$$\begin{aligned} \min_{\pi := \{\pi_{ij}\}} & \sum_{i=1}^N \sum_{j=1}^n \frac{1}{2} \|\mathbf{Y}_j - \mathfrak{G}_i\|^2 \pi_{ij} \\ \text{s.t.} & \sum_{j=1}^n \pi_{ij} = N^{-1}, \quad i \in \{1, 2, \dots, N\}, \\ & \sum_{i=1}^N \pi_{ij} = w_j^{(n)}(\mathbf{x}; \mathbf{X}^{(n)}), \quad j \in \{1, 2, \dots, n\}, \\ & \pi_{ij} \geq 0, \quad i \in \{1, 2, \dots, N\}, j \in \{1, 2, \dots, n\}. \end{aligned} \quad (3.3)$$

Here, any Nn -tuple $\pi := \{\pi_{ij} | i = 1, \dots, N, j = 1, \dots, n\}$ satisfying the constraints in (3.3) represents a transport plan from $U_d^{(N)}$ to $P_{w(\mathbf{x})}^{(n)}$ —that is, a discrete distribution over $\mathbb{R}^d \times \mathbb{R}^d$ with marginals $U_d^{(N)}$ and $P_{w(\mathbf{x})}^{(n)}$. Let

$$\pi^*(\mathbf{x}) = \{\pi_{ij}^*(\mathbf{x}) | i = 1, \dots, N, j = 1, \dots, n\}$$

be a solution of (3.3) (an optimal *transport plan*). Theorem 2.12(i) in Villani (2003) implies that its support $\text{supp}(\pi^*(\mathbf{x})) := \{(\mathfrak{G}_i, \mathbf{Y}_j) | \pi_{ij}^*(\mathbf{x}) > 0\}$ is *cyclically monotone*,⁸ hence, is contained in the graph of the *subdifferential* of a convex function. Therefore, the idea is to construct a smooth interpolation of $\pi^*(\mathbf{x})$ that maintains this property.

Note that, for any gridpoint $\mathfrak{G}_i, i \in \{1, \dots, N\}$, the constraints in (3.3) imply that there exists at least one $j \in \{1, \dots, n\}$ such that $(\mathfrak{G}_i, \mathbf{Y}_j) \in \text{supp}(\pi^*(\mathbf{x}))$. Since more than one such j may

exist, we choose the one which *gets the highest mass* from \mathfrak{G}_i and, in case of ties, the *smallest* one. Namely, let

$$\mathbf{T}^*(\mathfrak{G}_i | \mathbf{x}) := \arg \inf \left\{ \|\mathbf{y}\| : \mathbf{y} \in \text{conv}(\{\mathbf{Y}_J : J \in \arg \max_j \pi_{ij}^*(\mathbf{x})\}) \right\}, \quad (3.4)$$

where $\text{conv}(A)$ denotes the convex hull of a set $A \subset \mathbb{R}^d$. Since $\text{conv}(\{\mathbf{Y}_J : J \in \arg \max_j \pi_{ij}^*(\mathbf{x})\})$ is closed and convex in \mathbb{R}^d , (3.4) defines a unique $\mathbf{T}^*(\mathfrak{G}_i | \mathbf{x})$. Due to the cyclical monotonicity of $\text{supp}(\pi^*(\mathbf{x}))$, there exists a convex function $\varphi^*(\cdot | \mathbf{x}) : \mathbb{R}^d \rightarrow \mathbb{R}$ with subdifferential $\partial\varphi^*(\cdot | \mathbf{x})$ such that, for all

$$1 \leq i \leq N, \emptyset \neq \{\mathbf{Y}_j : (\mathfrak{G}_i, \mathbf{Y}_j) \in \text{supp}(\pi^*(\mathbf{x}))\} \subset \partial\varphi^*(\mathfrak{G}_i | \mathbf{x}).$$

Sub-differentials being convex, this entails

$$\text{conv}\{\mathbf{Y}_j : (\mathfrak{G}_i, \mathbf{Y}_j) \in \text{supp}(\pi^*(\mathbf{x}))\} \subset \partial\varphi^*(\mathfrak{G}_i | \mathbf{x}).$$

Consequently, $\{(\mathfrak{G}_i, \mathbf{T}^*(\mathfrak{G}_i | \mathbf{x})) : i = 1, \dots, N\}$ is cyclically monotone and satisfies the assumptions of Corollary 3.1 in Hallin et al. (2021). This implies the existence, for all \mathbf{x} , of a continuous cyclically monotone map $\mathbf{u} \mapsto \mathbf{Q}_{w, \pm}^{(n)}(\mathbf{u} | \mathbf{x})$, say, interpolating the N -tuple $(\mathfrak{G}_1, \mathbf{T}^*(\mathfrak{G}_1 | \mathbf{x})), \dots, (\mathfrak{G}_N, \mathbf{T}^*(\mathfrak{G}_N | \mathbf{x}))$, that is, such that $\mathbf{Q}_{w, \pm}^{(n)}(\mathfrak{G}_i | \mathbf{x}) = \mathbf{T}^*(\mathfrak{G}_i | \mathbf{x})$ for $i = 1, \dots, N$.

In particular, we proceed as in Hallin et al. (2021) by choosing the smooth cyclically monotone interpolation $\mathbf{u} \mapsto \mathbf{Q}_{w, \pm}^{(n)}(\mathbf{u} | \mathbf{x})$ with largest Lipschitz constant. Call *empirical conditional center-outward quantile function* of \mathbf{Y} given $\mathbf{X} = \mathbf{x}$ this continuous map from \mathbb{S}_d to \mathbb{R}^d . It defines the *empirical center-outward regression quantile regions and contours*

$$\begin{aligned} \mathbb{C}_{w, \pm}^{(n)}(\tau | \mathbf{x}) & := \mathbf{Q}_{w, \pm}^{(n)}(\tau \bar{\mathbb{S}}_d | \mathbf{x}) \\ \text{and } \mathbb{C}_{w, \pm}^{(n)}(\tau | \mathbf{x}) & := \mathbf{Q}_{w, \pm}^{(n)}(\tau \mathcal{S}_{d-1} | \mathbf{x}), \quad \tau \in (0, 1) \end{aligned} \quad (3.5)$$

which we are proposing as estimators of $\mathbb{C}_{\pm}(\tau | \mathbf{x})$ and $\mathbb{C}_{\pm}(\tau | \mathbf{x})$, respectively. The intersection $\bigcap_{\tau \in (0, 1)} \mathbb{C}_{w, \pm}^{(n)}(\tau | \mathbf{x})$ yields the *empirical conditional center-outward regression median region*. The definition

$$\mathbb{T}_{w, \pm}^{(n)}(\tau) := \left\{ (\mathbf{x}, \mathbf{Q}_{w, \pm}^{(n)}(\tau \bar{\mathbb{S}}_d | \mathbf{x})) | \mathbf{x} \in \mathbb{R}^m \right\}, \tau \in (0, 1)$$

of *empirical regression quantile tubes* naturally follows.

Remark 3.1. Note that the results of this section and the next one still hold for any continuous map with cyclically monotone graph satisfying

$$\begin{aligned} (\mathbf{u}_i, \mathbf{Q}_{w, \pm}^{(n)}(\mathbf{u}_i | \mathbf{x})) \in \text{conv}(\{\mathbf{Y}_j : (\mathbf{u}_i, \mathbf{Y}_j) \in \text{supp}(\pi^*(\mathbf{x}))\}) \\ \text{for all } i = 1, \dots, N. \end{aligned}$$

The reason for choosing the “smallest” \mathbf{y} in (3.4) is to have a “universal criterion.”

⁷We are dropping the superscripts (N) and (n) when no confusion is possible.

⁸Recall from Rockafellar (1970) that a set $S \subset \mathbb{R}^d \times \mathbb{R}^d$ is *cyclically monotone* if any finite subset $\{(x_{k_1}, y_{k_1}), \dots, (x_{k_v}, y_{k_v})\} \subset S, v \in \mathbb{N}$ is such that $\sum_{\ell=1}^{v-1} \langle y_{k_\ell}, x_{k_{\ell+1}} - x_{k_\ell} \rangle + \langle y_{k_v}, x_{k_1} - x_{k_v} \rangle \leq 0$, where $\langle \cdot, \cdot \rangle$ stands for the scalar product in \mathbb{R}^d .

3.2. Consistency

The objective of this section is to justify the definitions of Section 3.1 by showing the consistency of the empirical quantile regions and contours defined in (3.5). The asymptotic behavior of these regions and contours, quite naturally, depends on the regularity of the conditional distributions involved. In fact, as discussed before, when Assumption (R) does not hold, the population conditional quantile maps are not necessarily defined for every $\mathbf{u} \in \mathbb{S}_d$, but only for a subset of U_d -probability one. Consistency results can be obtained despite this a.s. definition provided that population quantile maps are extended into set-valued maps. The following theorem shows,⁹ under mild assumptions, that any possible limit of $\mathbf{Q}_{w,\pm}^{(n)}(\cdot | \mathbf{x})$ asymptotically belongs to the set $\mathbf{Q}_{\pm}(\cdot | \mathbf{X} = \mathbf{x})$.

Theorem 3.2. Let $(\mathbf{X}, \mathbf{Y}), (\mathbf{X}_1, \mathbf{Y}_1), \dots, (\mathbf{X}_n, \mathbf{Y}_n)$ be pairs of iid random vectors with values in $\mathbb{R}^m \times \mathbb{R}^d$ and let $w^{(n)}$ be a consistent sequence of weight functions. Then,

- (i) for every $\mathbf{u} \in \mathbb{S}_d$ and $\epsilon > 0$, $\mathbb{P}\left(\mathbf{Q}_{w,\pm}^{(n)}(\mathbf{u} | \mathbf{X}) \notin \mathbf{Q}_{\pm}(\mathbf{u} | \mathbf{X}) + \epsilon \mathbb{S}_d\right) \rightarrow 0$ as n and $N \rightarrow \infty$,
- (ii) for every $\tau \in (0, 1)$, $\mathbb{P}\left(\mathcal{C}_{\pm}^{(n)}(\tau | \mathbf{X}) \not\subset \mathcal{C}_{\pm}(\tau | \mathbf{X}) + \epsilon \mathbb{S}_d\right) \rightarrow 0$ as n and $N \rightarrow \infty$, and
- (iii) for every $\tau \in (0, 1)$, $\mathbb{P}\left(\mathcal{C}_{\pm}^{(n)}(\tau | \mathbf{X}) \not\subset \mathcal{C}_{\pm}^{(n)}(\tau | \mathbf{X}) + \epsilon \mathbb{S}_d\right) \rightarrow 0$ as n and $N \rightarrow \infty$.

Neater convergence results—avoiding the notion of set-valued maps—are obtained if it can be assumed that Assumption (R) holds, which implies that, for any $\mathbf{X} = \mathbf{x}$ and any $\mathbf{u} \in \mathbb{S}_d \setminus \{\mathbf{0}\}$, the set $\mathbf{Q}_{\pm}(\mathbf{u} | \mathbf{X} = \mathbf{x})$ is a singleton. Then, the map $\mathbf{u} \mapsto \mathbf{Q}_{\pm}(\mathbf{u} | \mathbf{X} = \mathbf{x})$ can be seen as continuous on $\mathbb{S}_d \setminus \{\mathbf{0}\}$, see Theorem 25.5 in Rockafellar (1970), hence, single-valued on $\mathbb{S}_d \setminus \{\mathbf{0}\}$ since the gradient of a convex function is single-valued at a point if and only if it is continuous at this point. We then can state the following theorem (see Appendix A for the proof), the second part of which describes the convergence of contours in terms of the Pompeiu-Hausdorff distance d_{∞} . Recall (Rockafellar and Wets 1998) that for two sets A and B in \mathbb{R}^d ,

$$d_{\infty}(A, B) := \inf\{v \geq 0 : A \subset B + v\mathbb{S}_d \text{ and } B \subset A + v\mathbb{S}_d\}.$$

Theorem 3.3. Let $(\mathbf{X}, \mathbf{Y}), (\mathbf{X}_1, \mathbf{Y}_1), \dots, (\mathbf{X}_n, \mathbf{Y}_n)$ be pairs of iid random vectors with values in $\mathbb{R}^m \times \mathbb{R}^d$ and let $w^{(n)}$ be a consistent sequence of weight functions. Suppose moreover that Assumption (R) holds. Then, for every compact $K \subset \mathbb{S}_d \setminus \{\mathbf{0}\}$, as n and $N \rightarrow \infty$,

$$\sup_{\mathbf{u} \in K} \left| \mathbf{Q}_{w,\pm}^{(n)}(\mathbf{u} | \mathbf{X}) - \mathbf{Q}_{\pm}(\mathbf{u} | \mathbf{X}) \right| \xrightarrow{\mathbb{P}} 0 \quad (3.6)$$

and, for every $\tau \in (0, 1)$ and $\epsilon > 0$,

$$\mathbb{P}\left(d_{\infty}\left(\mathcal{C}_{\pm}^{(n)}(\tau | \mathbf{X}), \mathcal{C}_{\pm}(\tau | \mathbf{X})\right) > \epsilon\right) \rightarrow 0. \quad (3.7)$$

Under the assumptions of Theorem 3.3, consistency in Pompeiu-Hausdorff distance of the quantile contours holds in case the median is a single point¹⁰—the continuity of quantile

maps then extends to the whole open unit ball. This, however, is not necessarily the case for $d > 3$ (Figalli 2018), and Pompeiu-Hausdorff consistency may fail due to the fact that our empirical version is continuous over \mathbb{S}_d while $\mathbf{Q}_{\pm}(\mathbf{0} | \mathbf{x})$ could be a set rather than a single point: convergence then holds along subsequences of $\mathbf{Q}_{w,\pm}^{(n)}(\cdot | \mathbf{x})$ to an element of $\mathbf{Q}_{\pm}(\mathbf{0} | \mathbf{x})$. This has an impact on convergence in terms of the Pompeiu-Hausdorff distance—although it does not affect the control over the asymptotic probability contents of quantile regions. More precisely, the following corollary holds (see Appendix A for the proof).

Corollary 3.4. Under the conditions of Theorem 3.3,

$$\mathbb{P}\left(\mathbf{Y} \in \mathcal{C}_{\pm}^{(n)}(\tau | \mathbf{X}) | \mathbf{X}\right) \xrightarrow{\mathbb{P}} \tau \text{ for all } \tau \in (0, 1)$$

as n and $N \rightarrow \infty$.

Under the weaker conditions of Theorem 3.2, in view of (A.12), some asymptotic control of the probability content of empirical regions is achieved. More precisely, letting $N = N(n)$ with $N(n) \rightarrow \infty$ as $n \rightarrow \infty$, for all $\tau \in (0, 1)$ and every subsequence $n_k \rightarrow \infty$, there exist further subsequences $n_{k_j} \rightarrow \infty$ such that $\limsup_{j \rightarrow \infty} \mathbb{P}\left(\mathcal{C}_{\pm}^{(n_{k_j})}(\tau | \mathbf{x}) | \mathbf{X} = \mathbf{x}\right) \leq \tau$ Lebesgue-a.e. in \mathbb{R}^m .

The above results, as well as the proposed regularization, are valid for any consistent sequence of weight functions. This, with adequate additional assumptions, includes most of the classical cases, such as the *kernel* and *classical nearest-neighbors* weight functions: see Appendix C. The classical k -nearest neighbors (C.2) are understood in the classical sense of the Euclidean distance in \mathbb{R}^m , which does not take into account the distribution $P_{\mathbf{X}}$ of \mathbf{X} . An alternative k -nearest neighbors weight function can be derived from a notion of nearness based on the ordering induced by empirical center-outward distribution functions.

This alternative weight function is obtained as follows. Fixing $\mathbf{x} \in \mathbb{R}^m$, first compute, as in Hallin et al. (2021), the empirical center-outward distribution function associated with

$$\frac{1}{n+1} \sum_{j=1}^n \delta_{\mathbf{X}_j} + \frac{1}{n+1} \delta_{\mathbf{x}} \in \mathcal{P}(\mathbb{R}^m).$$

That distribution function is the solution $T_{\mathbf{x}}^*$ of the minimization problem $\min_{T \in \Gamma_{n+1}} \sum_{k=0}^n |\mathbf{X}_k - T(\mathbf{X}_k)|^2$ where $\mathbf{X}_0 = \mathbf{x}$, Γ_{n+1} is the set of all bijections T between $\{\mathbf{x}, \mathbf{X}_1, \dots, \mathbf{X}_n\}$ and a *regular grid* $\mathfrak{G}^{(n+1)}$ of \mathbb{S}_m , of the form described in Section 3.1, consisting of $(n+1)$ gridpoints, denoted as $\mathfrak{G}_0, \mathfrak{G}_1, \dots, \mathfrak{G}_n$, obtained via a factorization of the form $n+1 = n_R n_S + n_0$ with $n_R, n_S, n_0 \in \mathbb{N}$ and $n_0 < \min(n_R, n_S)$. That grid then consists in the intersection of

- the n_S rays generated by an n_S -tuple $\mathbf{u}_1, \dots, \mathbf{u}_{n_S} \in \mathcal{S}_{m-1}$ of unit vectors such that $n_S^{-1} \sum_{j=1}^{n_S} \delta_{\mathbf{u}_j}$ converges weakly, as $n_S \rightarrow \infty$, to the uniform over \mathcal{S}_{m-1} and
- the n_R hyperspheres with center $\mathbf{0}$ and radii $j/(n_R + 1)$, $j = 1, \dots, n_R$,

⁹Note that, although N does not appear in the notation, $\mathbf{Q}_{w,\pm}^{(n)}$ depends on both N and n .

¹⁰This is always the case for $d = 2$ and $d = 3$: see Figalli (2018).

along with n_0 copies of the origin whenever $n_0 > 0$. Based on this grid, we define the sequence of discrete uniform measures

$$U_d^{(n+1)} := \frac{1}{n+1} \sum_{j=1}^{n+1} \delta_{\mathfrak{G}_j} \in \mathcal{P}(\mathbb{R}^m), N \in \mathbb{N}.$$

The resulting map T_x^* is defined only at the $(n+1)$ points $\mathbf{x}, \mathbf{X}_1, \dots, \mathbf{X}_n$, but, as in the previous section, it can be continuously extended (see also Hallin et al. 2021) to the whole space \mathbb{R}^m —call $\mathbf{F}_{\mathbf{x};\pm}^{(n)} : \mathbb{R}^m \rightarrow \mathbb{S}_m$ this extension—with the properties that $\mathbf{F}_{\mathbf{x};\pm}^{(n)}$ coincides with T_x^* on $\{\mathbf{x}, \mathbf{X}_1, \dots, \mathbf{X}_n\}$, is the gradient of a differentiable convex function with domain \mathbb{R}^m , and satisfies $\mathbf{F}_{\mathbf{x};\pm}^{(n)}(\mathbb{R}^m) \subset \overline{\mathbb{S}}_m$. We then define the set of k -nearest center-outward neighbors of \mathbf{x} as

$$\mathcal{K}_k^{(n)}(\mathbf{x}) := \{\mathbf{X}_j : \mathbf{F}_{\pm}^{(n)}(\mathbf{X}_j) \in \mathcal{N}_k(\mathbf{F}_{\mathbf{x};\pm}^{(n)}(\mathbf{x}))\}$$

where, for each $\mathbf{a} \in \mathbf{B}_m$ and $k \in \mathbb{N}$, $\mathcal{N}_k(\mathbf{a})$ denotes the set of k -nearest neighbors (in the sense of Euclidean distance) of \mathbf{a} . Define the center-outward nearest neighbor weight function as

$$w_j^{(n)}(\mathbf{x}; \mathbf{X}^{(n)}) := \frac{1}{k} \mathbb{1}_{\mathbf{X}_j \in \mathcal{K}_k^{(n)}(\mathbf{x})}, j = 1, \dots, k. \quad (3.8)$$

For a suitable choice of $k = k(n)$, center-outward nearest neighbors weight functions form a consistent sequence of weights (see Lemma 3.5, and Appendix A for a proof). Finally, proceed as in Section 3.1 with the estimation (3.3) of conditional quantile functions.

Lemma 3.5. If $k = k(n)$ is such that $k(n) \rightarrow \infty$ and $k(n)/n \rightarrow 0$ as $n \rightarrow \infty$, the sequence of weight functions defined in (3.8) is consistent in the sense of (3.2).

This means, in particular, that Theorem 3.2 applies when the weight function (3.8) is used under the assumptions of Lemma 3.5, and that the resulting estimators are consistent.

Theorems 3.2 and 3.3 provide *weak* (in probability) *consistency* results under minimal assumptions. For sequences of weights satisfying, as $n \rightarrow \infty$,

$$\sum_{j=1}^n w_j^{(n)}(\mathbf{X}; \mathbf{X}^{(n)}) Y_j \longrightarrow E[Y|\mathbf{X}] \text{ a.s.} \quad (3.9)$$

(*strongly consistent* sequences), the conclusion in Theorem 3.2 can be upgraded to *strong* (a.s.) *consistency*. For the particular case of k -nearest neighbors, (3.9) and a.s. consistency hold with (C.1) replaced by $\frac{k}{\log(n)} \rightarrow \infty$ and $k/n \rightarrow 0$, (Devroye 1982; Devroye et al. 1994).

4. Numerical Results

This section is devoted to a numerical assessment of the performance of the estimation procedures described in Section 3. We focus on the classical case of k -nearest neighbor weights. Numerical results about kernel weights can be found in Appendix D. We first analyze (Section 4.1) some simulated datasets—including the motivating example of Hallin et al. (2015)—then turn (Section 4.3) to real data. These examples showcase three important features of our estimators: their ability to capture heteroscedasticity, to deal (non-parametrically) with highly nonlinear regression, and to adapt to non-convex distributions and noise with \mathbf{x} -dependent distribution shapes.

4.1. Simulated Examples

4.1.1. Parabolic Trend, Periodic Heteroscedasticity, Spherical Densities

We start with the motivating example ($m = 1$, $d = 2$) considered in Hallin et al. (2015)

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X \\ X^2 \end{pmatrix} + \left(1 + \frac{3}{2} \sin\left(\frac{\pi}{2}X\right)^2\right) \mathbf{e}, \quad X \sim U_{[-2,2]} \\ \text{and } \mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{Id}) \quad (4.1)$$

where X and \mathbf{e} are mutually independent. The population conditional (on $X = x$) quantile contours are circles with radii depending on x and can be computed exactly; trend is parabolic, heteroscedasticity periodic.

Figures 3 and 4 illustrate the convergence of our estimated contours to their population counterparts. We show the conditional quantiles at eight different equispaced x -locations. Compared to Figure 1 in Hallin et al. (2015), our method produces slightly less smooth versions of the same contours, at least for smaller sample sizes. On the other hand, our method is able to capture non-convex contour shapes—something the method in Hallin et al. (2015) cannot, see Section 4.1.2. We also stress that our method is able to handle much larger datasets (the R package `modQR` used there cannot handle sample sizes over 10,000, as explained in the documentation). As shown in Figure 4, for sample sizes ($n = 500,000$) that can be handled with our method, empirical and population contours essentially coincide.

Model (4.1), as pointed out in Hallin et al. (2015), allows for testing the capacity of a method to estimate the trend while catching potential heteroscedasticity. A comparison with Figure 1 in Hallin et al. (2015) shows that both methods estimate the parabolic trend quite well, but that our method performs much better at capturing heteroscedasticity. Estimations are based on the classical k -nearest neighbors weights (see (C.2)), with $N = k$ selected by the cross-validation procedure described in Section E2 in the Appendix.

4.1.2. Nonlinear Trend, Periodic Heteroscedasticity, Nonspherical Densities

We now consider two models exhibiting the same nonlinear trend and heteroscedasticity as in Model (4.1), but with non-convex quantile contours. The first one is

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} \sin(\frac{2}{3}\pi X) \\ \cos(\frac{2}{3}\pi X) + X^2 + 2.645X^4 \end{pmatrix} + \begin{pmatrix} .575\sigma(X)e_1 \\ .25e_1^3 + \frac{\sigma(X)}{2.3}e_2 \end{pmatrix}, \quad (4.2)$$

where $X \sim U_{[-1,1]}$, $\mathbf{e} = (e_1, e_2)^T \sim \mathcal{N}(\mathbf{0}, \mathbf{Id})$ independent of X and $\sigma(X) = \sqrt{1 + \frac{3}{2} \sin(\frac{\pi}{2}X)^2}$. The second one is model (1.8), see Figure 2 in the Introduction.

Estimated medians and center-outward quantile contours are displayed in Figures 5 and 6, respectively, from simulated datasets of various sizes (top left, $n = 10,000$; top right, $n = 50,000$; bottom left, $n = 100,000$; bottom right, $n = 500,000$). For sample sizes $n = 10,000, 50,000$ and $100,000$, the number k of neighbors has been selected by the cross-validation procedure described in Section E.2 in the Appendix, while for $n = 500,000$ we chose $k = 14,400$. As before, we computed conditional quantiles at eight different equispaced x -locations. Conditional quantiles cannot be computed explicitly. However,

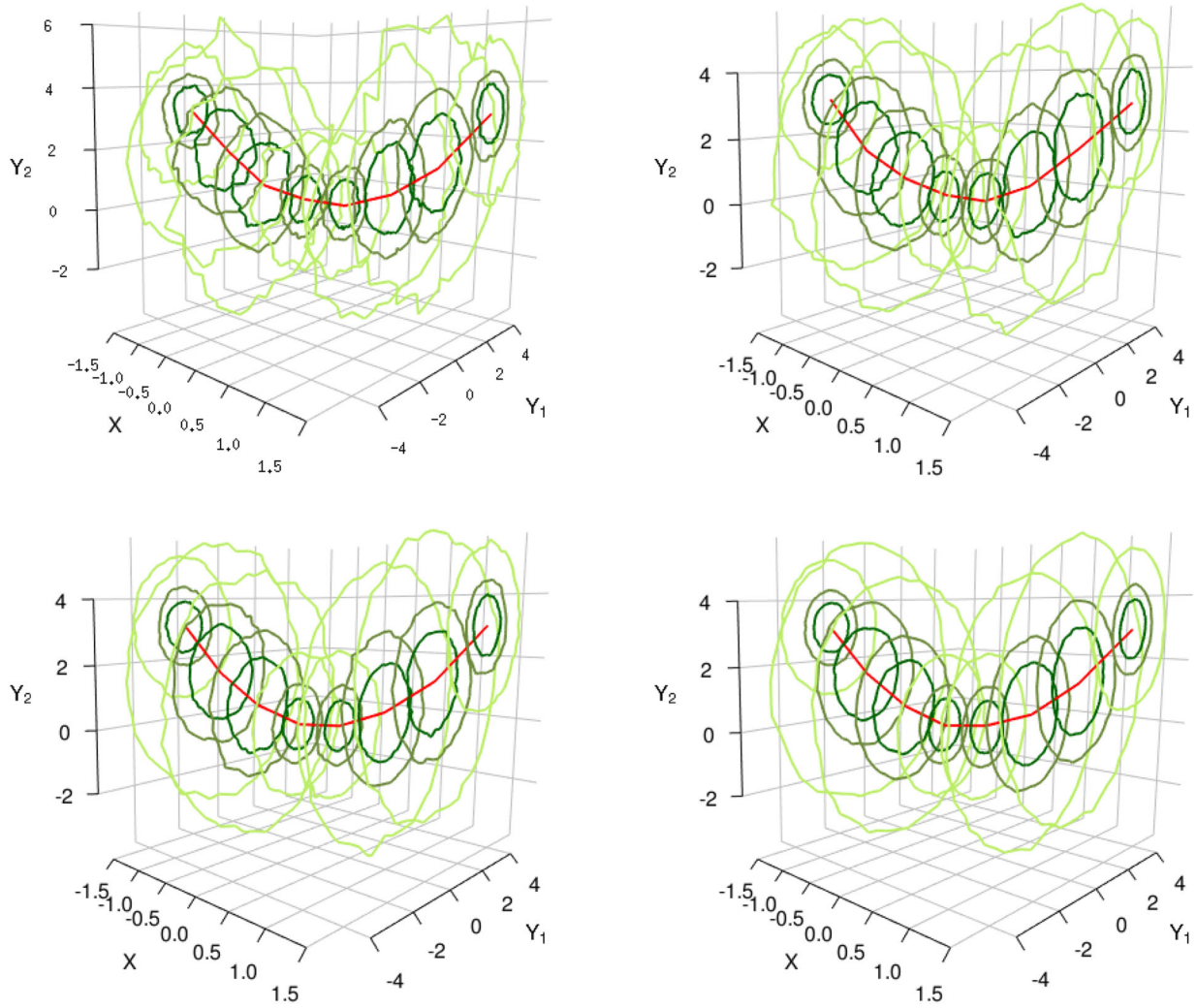


Figure 3. Estimated (sample sizes 10,000 in the upper left panel, 50,000 in the upper right panel, 100,000 in the lower left panel, and 500,000 in the lower right panel) quantile contours of order $\tau = 0.2$ (black), 0.4 (dark green), and 0.8 (light green) for Model (4.1); the (estimated) conditional center-outward medians are shown in red. Estimations are based on the classical k -nearest neighbors weights (see (C.2)) with $N = k$ selected by cross-validation.

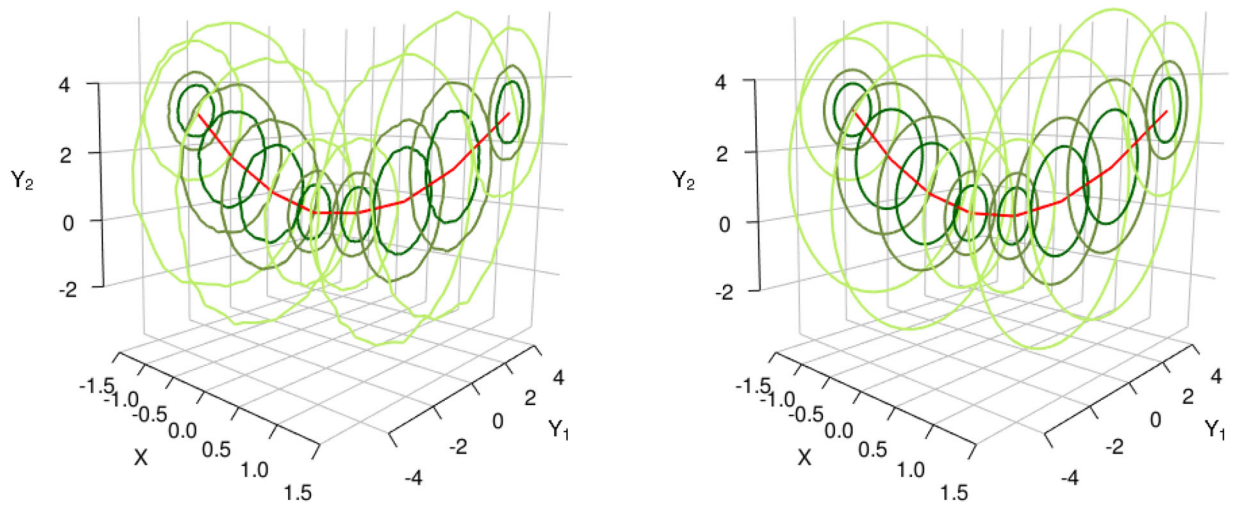


Figure 4. Estimated (sample size 500,000, left) versus population (right) quantile contours of order $\tau = 0.2$ (black), 0.4 (dark green), and 0.8 (light green) for Model (4.1).

as for model (4.1), we observe stabilization of the empirical contours to their population counterparts, which are quite accurately approximated. We also observe that the estimated medians

and contours provide a very good insight into the dependence of \mathbf{Y} on X , while the conditional center-outward quantile regions not only adapt to the nonlinear trend and heteroscedasticity

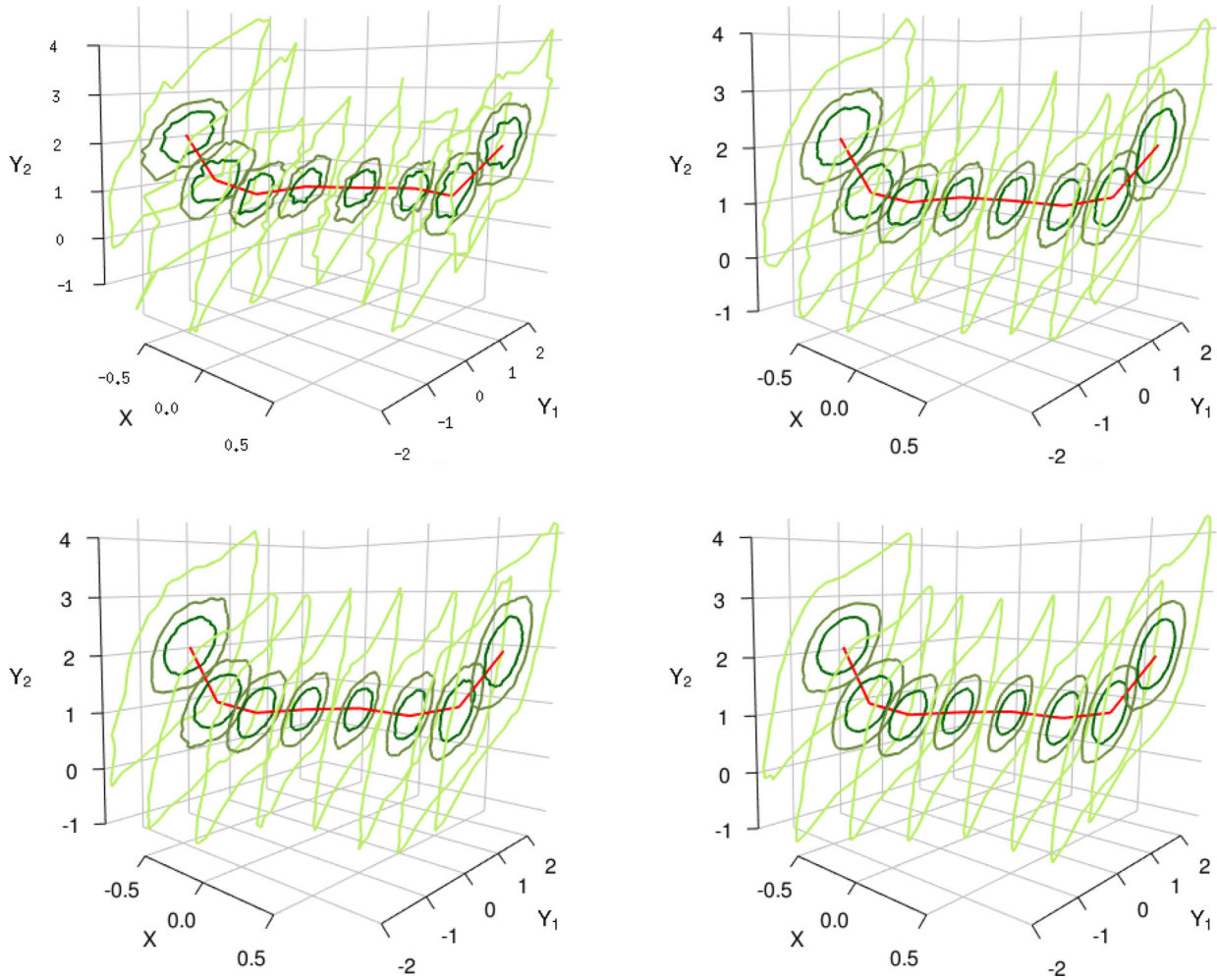


Figure 5. Empirical medians and empirical conditional center-outward quantile contours of order $\tau = 0.2, 0.4$, and 0.8 for model (4.2) and different sample sizes (top left, $n = 10,000$; top right, $n = 50,000$; bottom left, $n = 100,000$; bottom right, $n = 500,000$).

but also to the non-convexity of the conditional quantile regions.

4.1.3. Computational Issues

Computational costs and a cross-validation method for the selection of the number k of nearest neighbors are discussed in the online supplement (Sections E.1 and E.2).

4.2. Comparisons with Other Methods

Comparisons with existing methods are not easy and, to some extent, not entirely meaningful since the quantile concepts to be estimated differ. A comparison with the local bilinear halfspace-depth-based method of Hallin et al. (2015) is provided in Section 4.1. Below, we concentrate on the methods proposed by Carlier, Chernozhukov, and Galichon (2016) and Chakraborty (2003).

4.2.1. Comparison with Carlier, Chernozhukov, and Galichon (2016)

Carlier, Chernozhukov, and Galichon (2016)¹¹ introduced a linear multiple-output measure-transportation-based quantile regression model along with a numerical estimation procedure

for its practical computation, which we briefly outline. In our notation, their model takes the form

$$Q_{\pm}(\mathbf{u}|\mathbf{x}) = \boldsymbol{\beta}(\mathbf{u}) \cdot \mathbf{x} \quad (4.3)$$

where the value at fixed \mathbf{u} of the conditional quantile function is a linear function of \mathbf{x} (with a suitable choice of \mathbf{x} , (4.3) includes an intercept). Rather than the construction of conditional quantile regions, their objective is the estimation of $\boldsymbol{\beta}(\mathbf{u})$; their estimation method is based on a dual formulation of the original transportation problem (see (3.2)), the solution of which yields a potential function $\mathbf{u} \mapsto B(\mathbf{u})$ whose Jacobian is the function $\boldsymbol{\beta}$ to be estimated. In practice, this method yields an estimation of B over some grid of points from which an estimation of $\boldsymbol{\beta}$ follows via numerical differentiation. This works smoothly with uniformly equispaced grids over $[0, 1]^d$ providing good discretizations of the Lebesgue uniform over $[0, 1]^d$, but runs into difficulties with the spherical uniform over the unit ball. As a consequence, the method fails when center-outward quantiles are considered in (4.3).

As the objectives (estimation of $\boldsymbol{\beta}(\cdot)$ on the one hand, of $C_{\pm}(\cdot|\mathbf{x})$ on the other) are distinct, a full comparison between this approach and ours is not entirely meaningful—all the more so that their method does not apply under model (4.3). Denoting by $Q_{\square}(\cdot|\mathbf{x})$ the conditional quantile function associated with a

¹¹We thank Guillaume Carlier and his coauthors for kindly providing their codes.

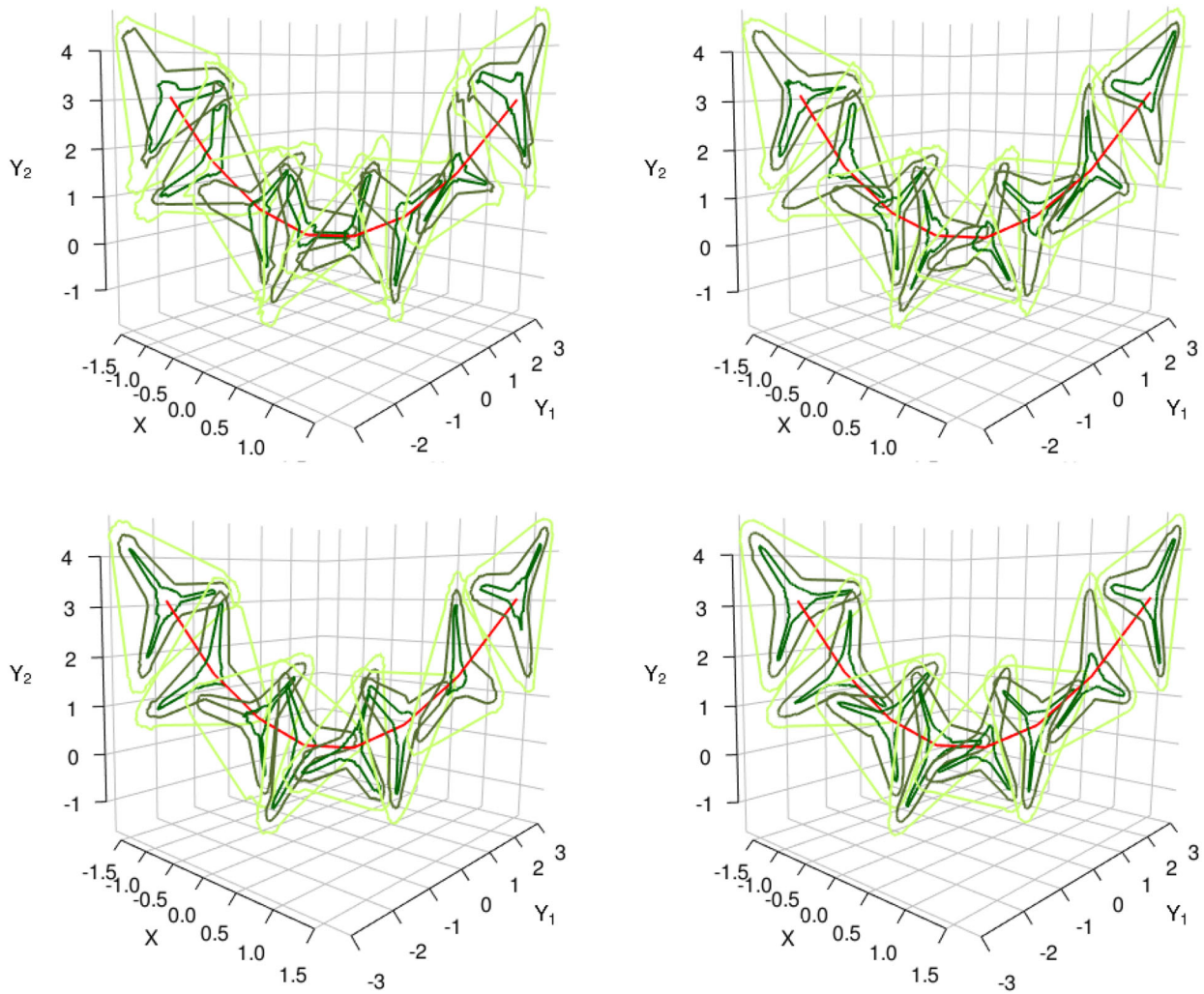


Figure 6. Empirical medians and empirical conditional center-outward quantile contours of order $\tau = 0.2, 0.4$, and 0.8 for model (1.8) and different sample sizes (top left, $n = 10,000$; top right, $n = 50,000$; bottom left, $n = 100,000$; bottom right, $n = 500,000$).

transport from the Lebesgue uniform over $[0, 1]^d$, a limited comparison is possible, though, between the estimators of the conditional medians $\beta((1/2, \dots, 1/2)|\mathbf{x}) \cdot \mathbf{x}$ obtained from running the Carlier, Chernozhukov, and Galichon (2016) algorithm and our conditional center-outward median $C_{\pm}(0|\mathbf{x})$, respectively.

Such a comparison has to be based on simulations from some data-generating process; we chose the model (under spherical \mathbf{e} , the population regression medians coincide)

$$(Y_1 \ Y_2)^{\top} = (X \ X)^{\top} + \sigma(X)\mathbf{e}^{\top}, \quad (4.4)$$

with $X \sim U(-2, 2)$, and $\sigma(X) = 1 + \frac{3}{2} \sin(\frac{\pi}{2}X)^2$, where $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{Id})$ is independent of X and model (4.1); the only difference between these two models is that (4.4) has a linear trend and satisfies (4.3) while (4.1), with its parabolic trend, does not. The linear Carlier, Chernozhukov, and Galichon (2016) estimation method and our nonparametric method were applied to samples from these two models ($n = 1000$ and $10,000$) to obtain estimates of $\beta((1/2, \dots, 1/2)|\mathbf{x}) \cdot \mathbf{x}$ and $C_{\pm}(0|\mathbf{x})$, respectively. The results are displayed in Figure 7. Not surprisingly, the linear estimator performs very well (essentially, it cannot be distinguished from the actual trend) when assumption (4.3) holds and rather badly in (4.1), where it completely fails to detect the parabolic trend. Our nonparametric

center-outward median estimator, on the contrary, performs quite well in both case (and extremely well for $n = 10,000$).

4.2.2. Comparison with Chakraborty (2003)

A totally different approach (Chakraborty 2003) to multiple-output quantile regression can be based on the geometric or spatial quantile concept introduced in Chaudhuri (1996). We refer to Koenen and Paindaveine (2022) for a recent discussion of this and related approaches. The value at $\mathbf{u} \in \mathbf{S}_d$ of the geometric quantile function of $\mathbf{Y} \sim P$ is defined as the minimizer of the loss function $q \mapsto \int_{\mathbb{R}^d} \{\|\mathbf{y} - q\| - \|\mathbf{y}\| - \mathbf{u} \cdot \mathbf{y}\} dP(\mathbf{y})$. When reindexed by their probability content τ , say, geometric quantile contours are interpreted as quantile contours of order τ . Chakraborty (2003) then defines the quantile function of P conditional on $\mathbf{X} = \mathbf{x}$ as the geometric quantile function of \mathbf{Y} conditional on $\mathbf{X} = \mathbf{x}$.

While the geometric quantile concept nicely extends the classical univariate L_1 definition of quantiles, it also presents some limitations: geometric quantiles are unbounded as $\|\mathbf{u}\| \rightarrow 1$ even for a compactly supported P and they fail (Hallin and Koenen 2024) to adapt even to elliptical contours. These drawbacks carry over to the conditional quantiles of Chakraborty

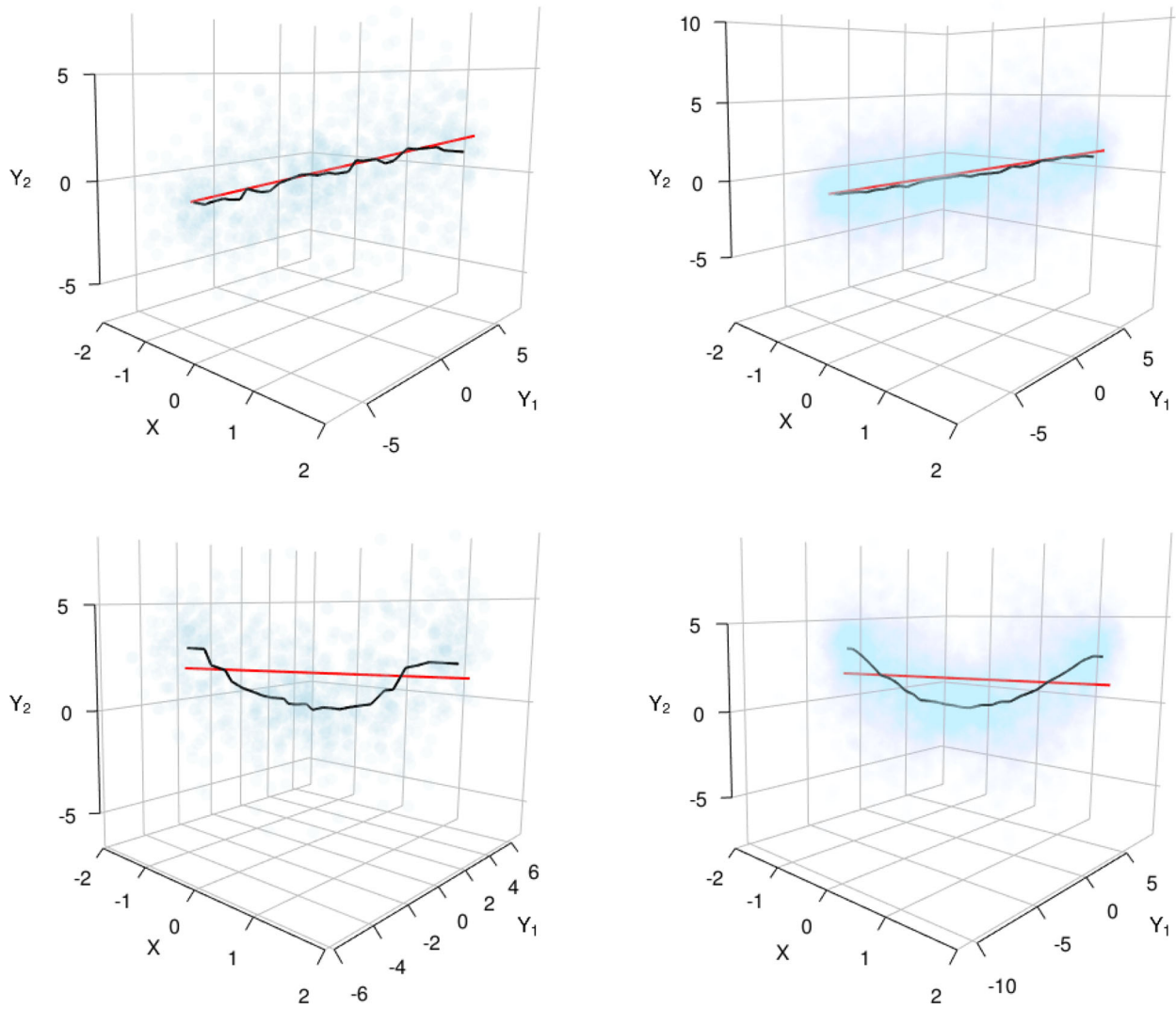


Figure 7. Linear (red) versus nonparametric (black) multiple-output conditional medians for models (4.4) (top) and (4.1) (bottom) for sample sizes $n = 1000$ (left) and $n = 10,000$ (right).

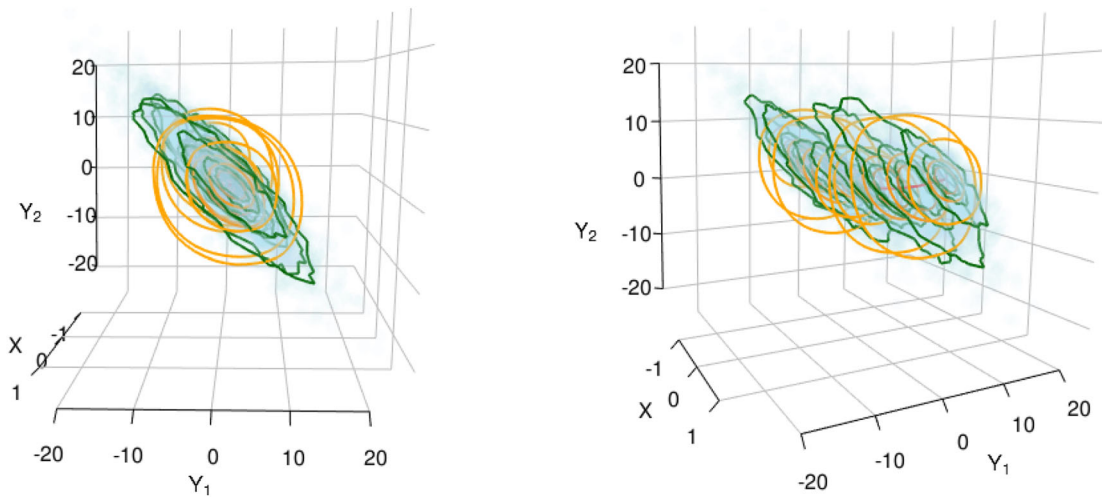


Figure 8. Conditional center-outward quantile contours (green) versus conditional geometric quantiles contours (orange) from a simulated sample ($n = 10,000$) of model (4.2); orders $\tau = 0.2, 0.4$, and 0.8 .

(2003). For the sake of comparison, we show in Figure 8 the conditional center-outward quantile contours (green) versus the conditional geometric quantiles (orange) of orders $\tau = 0.2, 0.4$,

and 0.8 evaluated from a simulated sample ($n = 10,000$) from model (4.2) (we thank Dimitri Konen for kindly providing his codes for the geometric quantiles). While the center-outward

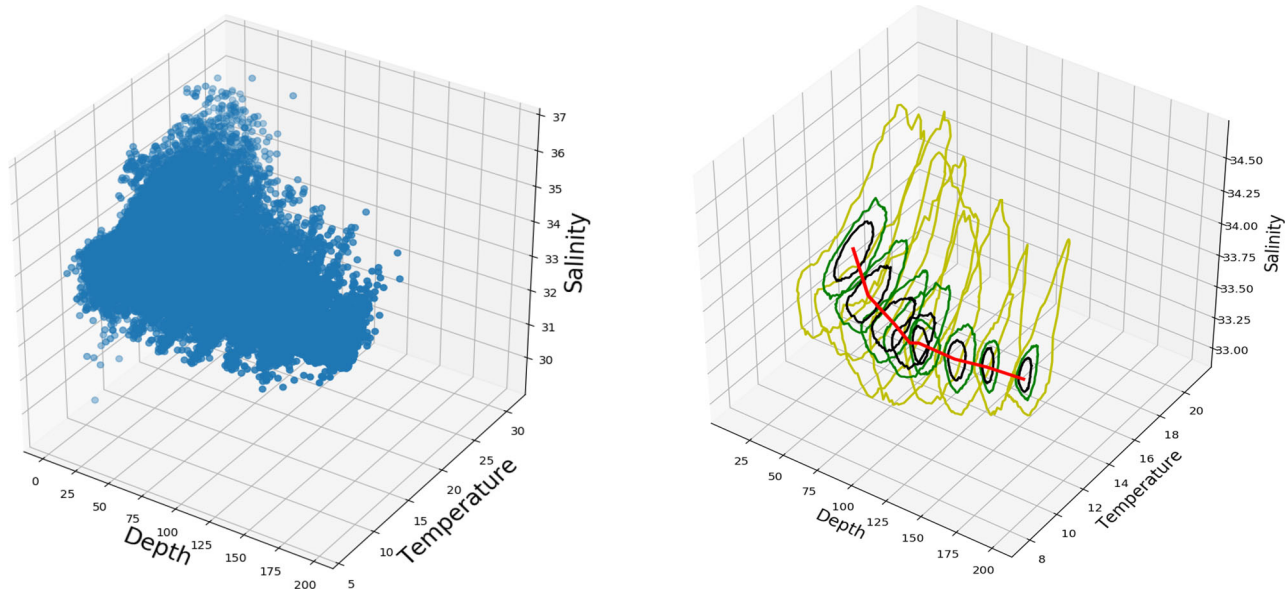


Figure 9. CalCOFI dataset. Left: the original dataset (depth, temperature, salinity) for $depth \leq 200$ (sample size $n = 505,829$). Right: the empirical conditional center-outward quantile contours of orders $\tau = 0.2$ (black), $\tau = 0.4$ (green), and $\tau = 0.8$ (yellow) and the empirical conditional center-outward median (red) for the multiple-output regression of $(Y_1, Y_2) = (temperature, salinity)$ with respect to $X = depth$. Estimation based on a k -nearest neighbors weight function with $k = 6401$.

contours nicely adapt to the shape of the data cloud, geometric quantile contours do not, and include large regions with very few observations.

4.3. Two Real-Data Examples

4.3.1. The CalCOFI Oceanographic Dataset

The dataset “CalCOFI Over 60 Years of Oceanographic Data,” available at <https://www.kaggle.com/sohier/calcofi>, contains the longest (1949-present) and most complete (more than 50,000 sampling stations; $n = 814,247$) time series of oceanographic data worldwide. Data collected at depths down to 500 meters include temperature, salinity, oxygen, phosphate, silicate, nitrate, and nitrite, chlorophyll, transmissometer, PAR, C14 primary productivity, phytoplankton biodiversity, zooplankton biomass, and zooplankton biodiversity. We are focusing here on the influence of $X = depth$ (meters) on the pair $\mathbf{Y} = (Y_1, Y_2)^T = (temperature, salinity)^T$ (degrees and grams of salt per kilogram, respectively).

Figure 9 shows the corresponding 3D observations, the estimated conditional center-outward quantile contours obtained from the same method as in Section 4.1 (nearest neighbors weight function with $k = 6401$) Figure F.1 in Appendix F (supplementary materials) shows the projections of the same contours on the $(depth, salinity)$ and $(depth, temperature)$ axes, respectively. Inspection of these figures reveals a nonlinear center-outward conditional median; heteroscedasticity also appears as the area of the conditional quantile regions clearly decreases as a function of depth, while a positive dependence between temperatures and salinity, which is present at the surface, gradually disappears as depth increases. The projection plots also provide clearer views on marginal dependencies. For example, the decrease of temperature as a function of depth is monotone and almost linear, while the dependence on depth of salinity is more complex, high at shallow depths, lower at medium depths, and higher again

at greater depths. However, these marginal analyses, to some degree, are hiding the heteroscedasticity effects (in particular, the dependence on depth of the relation between salinity and temperature) which are clearly visible in Figure 9. Since the dataset is quite large, we used a nearest neighbors weight function, see the comments in Section 4.1.1.

4.3.2. The Female ANSUR 2 Dataset

Our second example involves a smaller sample size n . The Female Anthropometric Survey of US Army Personnel (Female ANSUR 2 or Female ANSUR II) featured in this section consists in 93 direct measures and 41 derived ones, as well as three-dimensional head, foot, and whole-body scans of $n = 1986$ women of the U.S. army. These measurements were collected between October 4, 2010 and April 5, 2012, in May 2014 and May 2015; they are available online at <https://www.openlab.psu.edu/ansur2/>.

We want to analyze the relation between the covariate $X = stature$ (in centimeters) and the variable of interest $\mathbf{Y} = (foot\ length, tibial\ height)$ (both in centimeters). Figure 10 provides a 3D view of the center-outward quantile contours/tubes for levels $\tau = 0.2, 0.4$, and 0.7 along with the center-outward regression median (red) Figure F.2 in Appendix F (supplementary materials) provides the projections on the axes of the same contours. Inspecting these two figures reveals the absence of heteroscedasticity, the spherical shape of conditional distributions, and a roughly linear regression. Since the size of the model is not too large, a Gaussian kernel is convenient. The bandwidth was chosen as $h = 15$, which, up to scale changes, corresponds to $h = 0.2$ in Figure D.4.

5. Conclusions and Perspectives

Building on the concepts of center-outward quantiles recently developed in Hallin et al. (2021), we are proposing here a

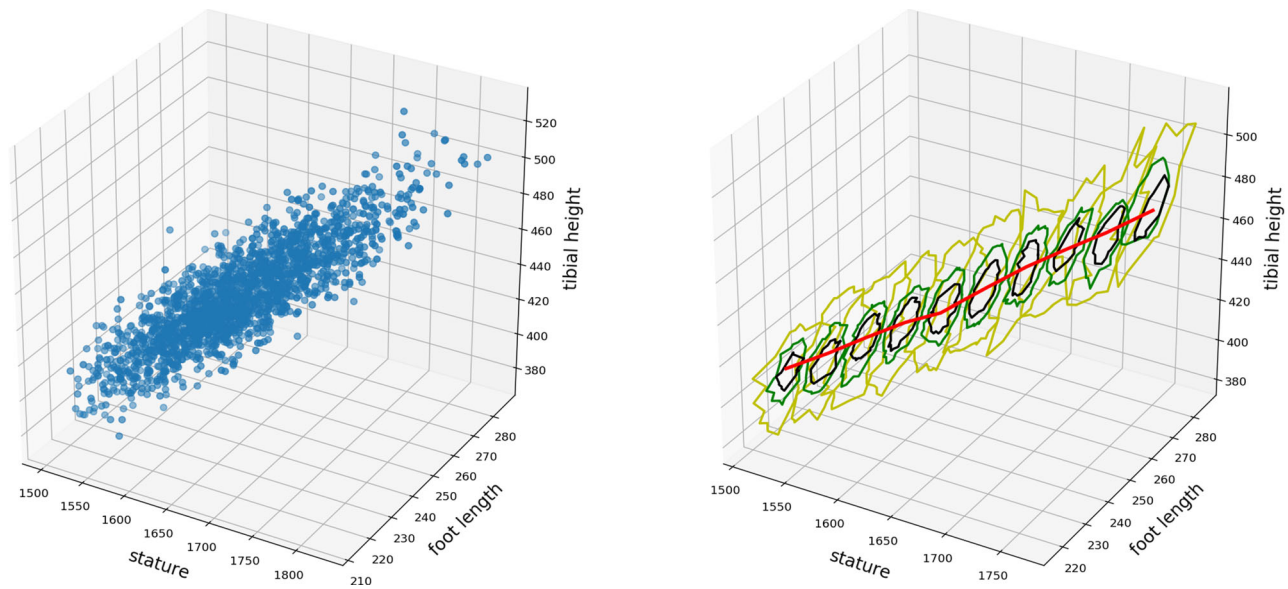


Figure 10. ANSUR 2 dataset (sample size 1986). Left: the original dataset of $X=\text{stature}$ and $(Y_1, Y_2)=(\text{foot length}, \text{tibial height})$. Right: the empirical conditional center-outward quantile contours for quantile orders $\tau = 0.2$ (black), 0.4 (green), and 0.7 (yellow) and the empirical conditional center-outward median (red) for the multiple-output regression of (Y_1, Y_2) with respect to X ; estimation based on a Gaussian kernel weight function with bandwidth $h = 15$.

fully nonparametric solution to the problem of multiple-output quantile regression. Contrary to earlier attempts, our solution is enjoying the quintessential property that the (conditional) probability content of its quantile regions is under control irrespective of the underlying distribution. This is only a first step into the multifarious applications of multiple-output quantile regression, though. Due to the minimality of the assumptions it requires, a completely agnostic nonparametric approach indeed is attractive, but also comes at a cost: linear or polynomial quantile regression remain justified whenever a priori knowledge of the analytical form of the regression is available and can be taken advantage of. A center-outward version of the results of Carlier, Chernozhukov, and Galichon (2016), thus, is highly desirable. Single-output quantile regression has been considered in a variety of contexts: survival analysis, longitudinal data, instrumental variable regression, directional, functional, and high-dimensional data, . . . Quantile regression versions of time-series models such as the quantile autoregressive model also have been investigated (Koenker and Xiao 2006). All these applications call for multiple-output extensions with important real-life consequences and should be based on the concept of center-outward quantile, regions, and contours. They are the subject of our ongoing research.

Supplementary Materials

Supplementary materials include seven appendices labeled A, B, . . . , G. Appendix A collects the proofs for Section 3. Appendix B provides the proofs for three lemmas (A.1–A.3) used in Appendix A. Appendix C describes the classical kernel and nearest-neighbors weight functions and Appendix D analyzes their respective performances. Appendix E discusses some of the computational issues of our estimation method: computation costs (E.1) and (E.2) a cross-validation approach to the choice of the number k of neighbors. Appendix F provides additional empirical results on the CalCOFI and ANSUR datasets. Finally Appendix G relates our approach to the current literature on optimal transport estimation.

Disclosure Statement

The authors report there are no competing interests to declare.

Funding

The research of Eustasio del Barrio is partially supported by PID2021-128314NB-I00, funded by MCIN/AEI/10.13039/501100011033/FEDER. The research of Alberto González-Sanz is partially supported by the AI Interdisciplinary Institute ANITI, which is funded by the French “Investing for the Future – PLA3” program under the Grant agreement ANR-19-PI3A-0004. Marc Hallin gratefully acknowledges the financial support of the Czech Science Foundation (grants GAČR22036365 and GA24-100788).

References

- Agarwal, G., Tu, W., Sun, Y., and Kong, L. (2022), “Flexible Quantile Contour Estimation for Multivariate Functional Data: Beyond Convexity,” *Computational Statistics and Data Analysis*, 168, 107400. [3]
- Breckling, J., and Chambers, R. (1988), “ m -Quantiles,” *Biometrika*, 75, 761–771. [2]
- Camehl, A., Fok, D., and Gruber, K. (2022), “Multivariate Quantile Regression Using Superlevel Sets of Conditional Densities,” Tinbergen Institute Discussion Papers 22-094/III, Tinbergen Institute. [2,3]
- Carlier, G., Chernozhukov, V., and Galichon, A. (2016), “Vector Quantile Regression: An Optimal Transport Approach,” *Annals of Statistics*, 44, 1165–1192. [3,4,10,11,14]
- Chakraborty, A., and Chaudhuri, P. (2014), “The Spatial Distribution in Infinite-Dimensional Spaces and Related Quantiles and Depths,” *Annals of Statistics*, 42, 1203–1231. [2]
- Chakraborty, B. (2003), “On Multivariate Quantile Regression,” *Journal of Statistical Planning and Inference*, 110, 109–132. [2,10,11,12]
- Chaudhuri, P. (1996), “On a Geometric Notion of Quantiles for Multivariate Data,” *Journal of the American Statistical Association*, 91, 862–872. [2,11]
- Chernozhukov, V., Galichon, A., Hallin, M., and Henry, M. (2017), “Monge-Kantorovich Depth, Quantiles, Ranks and Signs,” *Annals of Statistics*, 45, 223–256. [2]
- Chowdhury, J., and Chaudhuri, P. (2019), “Nonparametric Depth and Quantile Regression for Functional Data,” *Bernoulli*, 25, 395–423. [2]
- Deb, N., and Sen, B. (2023), “Multivariate Rank-based Distribution-Free Nonparametric Testing Using Measure Transportation,” *Journal of the American Statistical Association*, 118, 192–207. [2]

- del Barrio, E., González-Sanz, A., and Hallin, M. (2020), “A Note on the Regularity of Optimal-Transport-based Center-Outward Distribution and Quantile Functions,” *Journal of Multivariate Analysis*, 180, 104671. [5]
- Devroye, L. (1982), “Necessary and Sufficient Conditions for the Pointwise Convergence of Nearest Neighbor Regression Function Estimates,” *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 61, 467–481. [8]
- Devroye, L., Györfi, L., Krzyżak, A., and Lugosi, G. (1994), “On the Strong Universal Consistency of Nearest Neighbor Regression Function Estimates,” *Annals of Statistics*, 22, 1371–1385. [8]
- Figalli, A. (2018), “On the Continuity of Center-Outward Distribution and Quantile Functions,” *Nonlinear Analysis*, 177, 413–421. [5,7]
- Ghosal, P., and Sen, B. (2022), “Multivariate Ranks and Quantiles Using Optimal Transport: Consistency, Rates and Nonparametric Testing,” *Annals of Statistics*, 50, 1012–1037. [2]
- Girard, S., and Stupfler, G. (2017), “Intriguing Properties of Extreme Geometric Quantiles,” *REVSTAT*, 15, 107–139. [2]
- Hallin, M. (2022), “Measure Transportation and Statistical Decision Theory,” *Annual Review of Statistics and its Application*, 9, 401–424. [2]
- Hallin, M., and Konen, D. (2024), “Multivariate Quantiles: Geometric and Measure-Transportation-Based Contours”, in *Proceedings of the 17th International Conference of the Econometrics Society of Thailand*, Springer. Available at <https://arxiv.org/abs/2401.02499>. [2,11]
- Hallin, M., del Barrio, E., Cuesta-Albertos, J., and Matrán, C. (2021), “Distribution and Quantile Functions, Ranks and Signs in Dimension d : A Measure Transportation Approach,” *Annals of Statistics*, 49, 1139–1165. [2,6,7,8,13]
- Hallin, M., Hlubinka, D., and Hudecová, Š. (2023), “Efficient Fully Distribution-Free Center-Outward Rank Tests for Multiple-Output Regression and MANOVA,” *Journal of the American Statistical Association*, 118, 1923–1939. [2]
- Hallin, M., La Vecchia, D., and Liu, H. (2022), “Center-Outward Estimation for Semiparametric VARMA Models,” *Journal of the American Statistical Association*, 117, 925–938. [2]
- (2023), “Rank-based Testing for Semiparametric VAR Models: A Measure Transportation Approach,” *Bernoulli*, 29, 229–273. [2]
- Hallin, M., and Liu, H. (2023), “Center-Outward Rank- and Sign-based VARMA Portmanteau Tests: Chitturi, Hosking, and Li–McLeod Revisited,” *Econometrics & Statistics*, to appear. [2]
- Hallin, M., Lu, Z., Paindaveine, D., and Šiman, M. (2015), “Local Bilinear Multiple-Output Quantile/Depth Regression,” *Bernoulli*, 21, 1435–1466. [2,4,8,10]
- Hallin, M., and Mordant, G. (2023), “Center-Outward Multiple-Output Lorenz Curves and Gini Indices: A Measure Transportation Approach,” arXiv:2211.10822. [2]
- Hallin, M., Paindaveine, D., and Šiman, M. (2010), “Multivariate Quantiles and Multiple-Output Regression Quantiles: From L_1 Optimization to Halfspace Depth [with Comments and Rejoinder],” *Annals of Statistics*, 38, 635–703. [1]
- Hallin, M., and Šiman, M. (2018), “Multiple-Output Quantile Regression,” in *Handbook of Quantile Regression*, eds. R. Koenker, V. Chernozhukov, X. He, and L. Peng, pp. 185–207, Boca Raton, FL: CRC Press. [2]
- Koenker, R. (2005), *Quantile Regression*, Econometric Society Monographs, Cambridge: Cambridge University Press. [1]
- Koenker, R., and Bassett, G. (1978), “Regression Quantiles,” *Econometrica*, 46, 33–50. [1]
- Koenker, R., Chernozhukov, V., He, X., and Peng, L., eds. (2018), *Handbook of Quantile Regression*, Boca Raton, FL: CRC Press. [1]
- Koenker, R., and Xiao, Z. (2006), “Quantile Autoregression [with Comments and Rejoinder],” *Journal of the American Statistical Association*, 101, 980–1006. [14]
- Koltchinski, V. (1996), “M-Estimation and Spatial Quantiles,” in *Robust Statistics, Data Analysis, and Computer Intensive Methods: In Honor of Peter Huber’s 60th Birthday*, ed. H. Rieder, pp. 235–250. New York: Springer. [2]
- Koltchinski, V. (1997), “M-estimation, Convexity and Quantiles,” *Annals of Statistics*, 25, 435–477. [2]
- Konen, D. (2023), “Explicit Recovery of a Probability Measure from its Geometric Depth,” arXiv:2208.11551. [2]
- Konen, D., and Paindaveine, D. (2022), “Multivariate ρ -quantiles: A Spatial Approach,” *Bernoulli*, 28, 1912–1934. [1,11]
- Kong, L., and Mizera, I. (2012), “Quantile Tomography: Using Quantiles with Multivariate Data,” *Statistica Sinica*, 22, 1589–1610. [2]
- Lehmann, E., and Romano, J. (2005), *Testing Statistical Hypotheses*, Cham: Springer. [5]
- McCann, R. J. (1995), “Existence and Uniqueness of Monotone Measure-Preserving Maps,” *Duke Mathematical Journal*, 80, 309–323. [2]
- Merlo, L., Petrella, L., Salvati, N., and Tzavidis, N. (2022), “Marginal M -quantile Regression for Multivariate Dependent Data,” *Computational Statistics & Data Analysis*, 173, 107500. [2]
- Nagy, S. (2021), “Halfspace Depth Does Not Characterize Probability Distributions,” *Statistical Papers*, 62, 1135–1139. [2]
- Paindaveine, D., and Šiman, M. (2011), “On Directional Multiple-Output Quantile Regression,” *Journal of Multivariate Analysis*, 102, 193–212. [2]
- (2012), “Computing Multiple-Output Regression Quantile Regions from Projection Quantiles,” *Computational Statistics*, 27, 29–49. [2]
- Peyré, G., and Cuturi, M. (2019), “Computational Optimal Transport with Applications to Data Science,” *Foundations and Trends® in Machine Learning*, 11, 355–607. [6]
- Rockafellar, R., and Wets, R. J.-B. (1998), *Variational Analysis*, New York: Springer. [5,7]
- Rockafellar, R. T. (1970), *Convex Analysis*, Princeton Mathematical Series. Princeton, NJ: Princeton University Press. [5,6,7]
- Serfling, R. (2002), “Quantile Functions for Multivariate Analysis: Approaches and Applications,” *Statistica Neerlandica*, 56, 214–232. [1]
- (2019), “Depth Functions on General Data Spaces. I: Perspectives, with Consideration of “Density” and “Local” Depths. II: Formulation and Maximality, with Consideration of the Tukey, Projection, Spatial, and “Contour” Depths,” <https://www.utdallas.edu/~serfling>. [1]
- Serfling, R., and Zuo, Y. (2000), “General Notions of Statistical Depth Function,” *Annals of Statistics*, 28, 461–482. [1]
- Shi, H., Drton, M., Hallin, M., and Han, F. (2021), “On Universally Consistent and Fully Distribution-Free Rank Tests of Vector Independence,” *Annals of Statistics*, 50, 1933–1959. [2]
- Shi, H., Drton, M., Hallin, M., and Han, F. (2024), “Center-Outward Sign- and Rank-based Quadrant, Spearman, and Kendall Tests for Multivariate Independence,” *Bernoulli*, to appear. [2]
- Stone, C. J. (1977), “Consistent Nonparametric Regression,” *Annals of Statistics*, 5, 595–620. [1,5]
- Tukey, J. W. (1975), “Mathematics and the Picturing of Data,” *Proceedings of the International Congress of Mathematicians, 1975*, 2, 523–531. [1]
- Villani, C. (2003), *Topics in Optimal Transportation*, Providence, RI: American Mathematical Society. [6]
- Wei, Y. (2008), “An Approach to Multivariate Covariate-Dependent Quantile Contours with Application to Bivariate Conditional Growth Charts,” *Journal of the American Statistical Association*, 103, 397–409. [2]