REVSTAT – Statistical Journal
Volume 18, Number 3, July 2020, 257–280

# PARAMETRIC ELLIPTICAL REGRESSION QUANTILES

Authors: Daniel Hlubinka
– Department of Probability and Mathematical Statistics,
Faculty of Mathematics and Physics, Charles University,
Czech Republic
hlubinka@karlin.mff.cuni.cz

Miroslav Šiman
– The Czech Academy of Sciences,
Institute of Information Theory and Automation,
Czech Republic

Received: February 2017 Revised: November 2017 Accepted: December 2017

Abstract:

- The article extends linear and nonlinear quantile regression to the case of vector responses by generalizing multivariate elliptical quantiles to a regression context. In particular, it introduces parametric elliptical quantile regression in a general nonlinear multivariate heteroscedastic framework and discusses, investigates, and illustrates the new method in some detail, including basic properties, various parametrizations, possible heteroscedastic patterns, related computational issues, model validation, and a real biometric data example. The method seems suitable for multiresponse regression models with symmetric errors, especially if the dimension of responses is less than ten and if the right parametrization of the model follows from the context.

Key-Words:

- *multiple-output regression; quantile regression; nonlinear regression; elliptical quantile.*

AMS Subject Classification:

- 62H12, 62J02, 62G05, 62G15, 62G35.

## 1. INTRODUCTION

Quantiles fully describe univariate probability distributions and may be very useful for statistical inference. Scalar random variables and their quantiles can often be expected to depend on some influential factors whose precise impact can be analyzed in the quantile regression framework, introduced in [23] and surveyed in [22]. Under weak moment assumptions, it models the entire conditional distribution of interest and not only its mean as the least squares approach. Therefore, it can reliably reveal even subtle changes in the conditional distribution that usually remain hidden in a conventional statistical analysis despite their possibly very important consequences. In fact, it is the tails of such conditional distributions that often contain much useful information and are thus very interesting for researchers in various fields such as finance and insurance, meteorology and climatology, labor and public economics, reliability and quality management, developmental studies, and medicine.

The everyday reality is usually intrinsically multivariate, and its successful analysis thus asks for multivariate quantiles. Unfortunately, they cannot be defined in a universally acceptable way because there exists no canonical way of ordering multivariate points and because all the attractive properties of univariate quantiles cannot be met simultaneously in a single multivariate quantile concept. Consequently, there already exist dozens of different multivariate quantile proposals that are usually based on data depth or spatial ranks, norm minimization or M-estimation, inversion of mappings, gradients, or generalized quantile processes; see, e.g., [33] for an overview.

Despite the abundance of the literature on multivariate quantiles (also called location quantiles), their regression generalizations are still scarce; see [18]. They may be either parametric (when the overall regression dependence is supposed to have a particular functional form), or nonparametric (when the overall dependence pattern is unknown). In the latter case, it is often possible to assume that the regression dependence is locally polynomial, which opens the door to the spline or locally polynomial (or, kernel) approach. Therefore, it makes perfect sense to call multivariate regression quantiles after the way they were obtained as parametric, nonparametric, or locally polynomial, for example. On the one hand, the parametric regression approach requires relatively strong assumptions regarding the particular form of regression dependence, on the other hand, it allows for general designs and implies standard consistency rates of related estimators (unlike its nonparametric competitors).

Most of the existing definitions of multivariate regression quantiles follow a directional strategy. They first define directional regression quantiles as simple objects (typically points or hyperplanes) and then use the directional objects for all directions to construct the resulting multivariate regression quantile (contour or region). The promising parametric proposals presented in [15, 16] and [29] are quite representative of this category and lead to the same multivariate regression quantile regions. Therefore, they will be considered as an established parametric golden standard and used as a benchmark hereinafter. They define a polyhedral multivariate regression quantile as the intersection of all directional regression quantile halfspaces of the same quantile level. They are implementable by means of [30, 31] and [2, 3], and applied, e.g., in [34] and [35]. The other proposals with directional flavor include [8], [25], [6], [9], [37], [26], [14], [7] and [4].

The alternative approach is not directional but direct (or, global) because it defines multivariate regression quantiles and related contours and regions directly, i.e., without any auxiliary directional construction. Apart from the very recent (but not affine equivariant) proposal of [5] inspired by [10], this category mainly includes various regression extensions of the two proposals of multivariate quantiles with elliptical shapes (or, elliptical quantiles) that were presented in [20] and [21]. The former proposal was motivated by linear quantile regression, included even a heuristic definition of locally constant elliptical quantiles, and employed only convex optimization that turned out very useful for its analysis. Unfortunately, it could not be extended within its convex optimization setting to include robust or flexible parametric regression quantiles, which is why the latter generalized multivariate elliptical quantile concept was proposed as a remedy in [21]. It could not rely on convex optimization any more, but, on the other hand, it was very general and even covered the former approach as a special case, after a suitable reparametrization.

Now the parametric regression extension of the generalized multivariate elliptical quantiles of [21] is discussed, investigated, and illustrated here in a very general nonlinear heteroscedastic framework. An important particular case with unique features has been briefly introduced in [17] together with its examination by means of convex analysis. It is nicely complemented with the general theory derived in this article.

It should also be mentioned for the sake of completeness, that the generalized parametric elliptical regression quantiles considered here bear some similarity to multivariate regression S-estimators and their modifications (see, e.g., [1], [36], and [32]) that are not used for defining multivariate regression quantiles but also result from some location-scale or regression-scale models where the determinant of the shape-defining matrix plays a crucial role.

As the parametric elliptical regression quantiles also roughly order the regression space, they remotely resemble the depth-like notions for regression observations (see, e.g., [15], [29], [34], and [14]).

In [21], the generalized multivariate elliptical quantiles have been shown useful for symmetric distributions and highly competitive with the benchmark introduced above for elliptical distributions. Their parametric regression extensions appear to preserve most of their properties, but they should also be used only if their conditionally elliptical shape is acceptable and, ideally, if the conditional distribution is at least centrally symmetric, which is fortunately the case of all widely used error distributions. Then they are roughly on par with the benchmark in terms of natural nestedness, equivariance properties, and the ability to change with the quantile level and to capture the symmetry or ellipticity of the underlying conditional distribution.

However, the generalized parametric elliptical regression quantiles then also excel in other important aspects. Indeed, unlike the benchmark,

- **(1)** they can easily incorporate homoscedasticity and many other types of a priori information regarding their conditional scales, shapes, and centers,
- **(2)** their quantile levels can correspond directly to their probability content,
- **(3)** they can be parametrized flexibly and very naturally by means of their conditional centers, shape matrices, and inflation (scaling) factors (whose estimates seem very useful for goodness-of-fit tests or for statistical inference regarding conditional location, dispersion, symmetry, or ellipticity),
- **(4)** they can be quite robust to outliers,

(**5**) they can work well even in complicated cases involving nonlinear trend or heteroscedasticity, and

(**6**) their computation can be feasible even in the sample cases involving moderate dimensions and large data sets.

In fact, their development seems motivated by the lack of a multivariate regression quantile concept with such a combination of favorable properties. Of course, some of them hold only under certain assumptions on the joint distribution of responses and regressors and on the parametrization of the model. Nevertheless, (1), (3), and (6) are totally out of reach of any directional multivariate quantile regression method.

Most of the following text only clarifies and demonstrates the vaguely stated properties of the generalized elliptical regression quantiles (and the conditions of their validity). As they generally do not result from convex optimization, their computation in the sample case may be quite complicated and their uniqueness may not be guaranteed. Nevertheless, they must be unique in certain special cases including those of [17], and such a possible ambiguity is common to many popular robust or nonlinear estimators. It might even be viewed as a positive feature in some cases involving multimodal conditional distributions that may arise easily in the context of mixtures; see, e.g., [12]. Until the uniqueness issues are satisfactorily resolved, it is nevertheless recommended to use the generalized parametric elliptical regression quantiles cautiously, to experiment with various initial values for their computation in the sample case, and to prefer linearity in their parametrization whenever possible.

Although the generalized parametric elliptical regression quantiles presented here are still somewhat rigid due to the ellipticity woven into their definition, they are definitely worthy of wide attention and careful investigation because there is apparently no other multivariate quantile regression methodology enabling joint parametric nonlinear modeling of both trend and heteroscedasticity without any specific distributional assumptions. It seems that the parametric elliptical quantile regression presented here has great potential and that it could be used with benefits for vector responses in the same fields as the univariate quantile regression or wherever else the whole conditional response distributions or their tails or covariance structures are of interest. That is to say that (various) multivariate regression quantiles have already proved very useful in several instances, e.g., in investigating the dependence

(**1**) of a few kinds of expenditures on the total income [5],

(**2**) of both systolic and diastolic blood pressures on age [6] or on age and BMI [9],

(**3**) of sales growth and sales profitability on the creativity test score in evaluating the performance of salespersons [6],

(**4**) of weight and height on age [37, 26],

(**5**) of a few product characteristics on the time of production to take the tool wear into consideration in the definition of a precision index [35],

(**6**) of length/height or weight and head circumference on age [27],

(**7**) of female thigh and calf maximum girths on age, height, weight or BMI [15, 14],

(**8**) of male life expectancy and death rate on the GNP per capita [29], or

(**9**) of a few financial time series [11, 4].

Some of the cited articles describe the application and its benefits in detail and should be consulted in case of any remaining doubts.

This article further proceeds as follows. Section 2 presents necessary notation and introduces the definition of generalized elliptical regression quantiles, Section 3 studies their basic properties in the population case, Section 4 discusses their parametrization, Section 5 uses them to classify multivariate heteroscedasticity, Section 6 deals with their computation in the sample case, Section 7 proposes some tools for their validation, Section 8 illustrates them with a few carefully designed demo examples, Section 9 applies them to a referential biometric dataset, and concluding Section 10 comments on the previous results and achievements. Applied statisticians reading the article for the first time may skip the text after Definition 2.1 and go directly to Section 4 or 8.

## 2. DEFINITIONS AND NOTATION

Consider a general regression setup where an $m$-variate stochastic vector of responses $\boldsymbol{Y} = \big(Y^{(1)}, ..., Y^{(m)}\big)' \in \mathbb{R}^m$ is to be explained with the aid of the corresponding $p$-variate regressor $\boldsymbol{Z} \in \mathbb{R}^p$, and $(\boldsymbol{Y}', \boldsymbol{Z}')'$ has an absolutely continuous distribution with a density differentiable almost everywhere.

Recall that the standard location and regression quantiles of [23] can be defined for any $\tau \in (0,1)$ by means of the non-negative convex real-valued check function $\rho_\tau(t) = t\big(\tau - \mathrm{I}(t<0)\big)$ $= \max\big\{(\tau-1)t,\, \tau t\big\}$ with a unique minimum. This function was also used in [20, 21] for defining two types of location elliptical quantiles. Here the second proposal is extended to a general parametric regression setup.

The next definition is rather complicated because it deals with the whole class of parametric elliptical regression quantiles indexed by quantile levels ($\tau$) and certain monotone functions ($g$), and because the natural parameters characterizing the shape of possible elliptical regression quantile contours ($\varepsilon_{g,\tau}$) themselves depend on a common parameter vector ($\theta$). Only its optimal value ($\theta_\tau$) resulting from a minimization problem is used in the definition.

**Definition 2.1.** For any $\tau \in (0,1)$ and any function $g$ specified below, the parametric elliptical regression $\tau$-$g$-quantile (contour) $\varepsilon_{g,\tau}(\boldsymbol{Y},\boldsymbol{Z})$ and the corresponding lower and upper parametric $\tau$-$g$-quantile regression regions $\mathcal{E}^-_{g,\tau}(\boldsymbol{Y},\boldsymbol{Z})$ and $\mathcal{E}^+_{g,\tau}(\boldsymbol{Y},\boldsymbol{Z})$ can be defined by means of the shape (matrix), trend (vector), and scale (scalar) quantile parameters $\mathbb{A}_\tau(\boldsymbol{\theta},\boldsymbol{z}) \in \mathbb{R}^{m\times m}$, $\boldsymbol{s}_\tau(\boldsymbol{\theta},\boldsymbol{z}) \in \mathbb{R}^m$, and $c_\tau(\boldsymbol{\theta},\boldsymbol{z}) \in \mathbb{R}$ depending on $\boldsymbol{z} \in \mathbb{R}^p$ as well as on a common parameter vector $\boldsymbol{\theta} = (\theta_1, ..., \theta_q)' \in \mathbb{R}^q$:

$$\begin{aligned}
\varepsilon_{g,\tau}(\boldsymbol{Y},\boldsymbol{Z}) &= \Big\{(\boldsymbol{y},\boldsymbol{z}) \in \mathbb{R}^{m+p} \colon\, h_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{y},\boldsymbol{z}) = 0\Big\}\,,\\
\mathcal{E}^-_{g,\tau}(\boldsymbol{Y},\boldsymbol{Z}) &= \Big\{(\boldsymbol{y},\boldsymbol{z}) \in \mathbb{R}^{m+p} \colon\, h_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{y},\boldsymbol{z}) < 0\Big\}\,,\\
\mathcal{E}^+_{g,\tau}(\boldsymbol{Y},\boldsymbol{Z}) &= \Big\{(\boldsymbol{y},\boldsymbol{z}) \in \mathbb{R}^{m+p} \colon\, h_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{y},\boldsymbol{z}) \ge 0\Big\}\,,
\end{aligned}$$

where

$$h_\tau(\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{z}) = g\Big(\big(\boldsymbol{y} - \boldsymbol{s}_\tau(\boldsymbol{\theta},\boldsymbol{z})\big)'\, \mathbb{A}_\tau(\boldsymbol{\theta},\boldsymbol{z})\, \big(\boldsymbol{y} - \boldsymbol{s}_\tau(\boldsymbol{\theta},\boldsymbol{z})\big)\Big) - c_\tau(\boldsymbol{\theta},\boldsymbol{z})\,,$$

$g(t)\colon [0,\infty) \mapsto [0,\infty)$ is a suitable strictly increasing smooth function such that $g(0) = 0$, and $\boldsymbol{\theta}_\tau$ minimizes the objective function

$$\text{(OF)} \qquad \Psi_\tau(\boldsymbol{\theta}) = \mathrm{E}\,\rho_\tau\big(h_\tau(\boldsymbol{\theta},\boldsymbol{Y},\boldsymbol{Z})\big)$$

over the whole parametric space $\Theta_\tau \subset \mathbb{R}^q$, $\overline{\Theta}_\tau = \overline{\Theta^\circ_\tau}$, subject to a regularity constraint on $\mathbb{A}_\tau$ ensuring that $\mathbb{A}_\tau(\boldsymbol{\theta}, \boldsymbol{z}) \in \mathbb{R}^{m\times m}$ is always symmetric positive definite (its choice is discussed below). The definition also tacitly assumes that the expectation in (OF) is finite and that its partial derivatives with respect to $\boldsymbol{\theta}$ are exchangeable with the expectation sign.

The sets $\varepsilon_{g,\tau}(\boldsymbol{Y}, \boldsymbol{Z}) \cap \{(\boldsymbol{y}, \boldsymbol{z}) \in \mathbb{R}^{m+p} \colon \boldsymbol{z} = \boldsymbol{z}_0\}$, defined for any fixed $\boldsymbol{z}_0 \in \mathbb{R}^p$, will be conveniently called elliptical $\tau$-$g$-quantile $\boldsymbol{z}_0$-cuts.

As far as the terminology is concerned, all the quantile-related adjectives, prefixes, indices, and arguments may be omitted on condition that they are either clear from the context or irrelevant to the statement being made.

Note that all the regression $\tau$-$g$-quantile $\boldsymbol{z}_0$-cuts are ellipsoids and that their definition resembles that of multivariate elliptical quantiles of [21] if $\mathbb{A}_\tau$, $\boldsymbol{s}_\tau$, and $c_\tau$ are independent of $\boldsymbol{z}$ and the regularity constraint is of the form $\det \mathbb{A}_\tau(\boldsymbol{\theta}, \boldsymbol{z}) = 1$. This constraint seems optimal for achieving the best possible equivariance properties of the resulting elliptical regression $\tau$-quantile entities and also from the statistical point of view, see [28], which is why it is exclusively considered here. This does not necessarily imply complete uselessness of all the other possible regularizations based on the eigenvalues of either $\mathbb{A}_\tau$ itself or of its product with a positive regressor-dependent scale factor; see [20] for some alternatives.

The definition of multivariate elliptical regression $\tau$-$g$-quantiles is obviously very general. First of all, it allows for very general trend and heteroscedastic patterns with possible nonlinearity in unknown parameters and with arbitrary $\tau$-dependence of $g$, $q$, $\Theta_\tau$, and the specifications for $\mathbb{A}_\tau$, $\boldsymbol{s}_\tau$, and $c_\tau$. It also permits quite general interdependencies between $\mathbb{A}_\tau$, $\boldsymbol{s}_\tau$, and $c_\tau$ thanks to their common dependence on the same parametric vector. Nevertheless, it is recommended that practitioners invoke simplicity and linearity whenever possible and reduce the use of interdependencies to the absolute minimum.

Of course, if there is any information regarding $\boldsymbol{\theta}_\tau$ available in advance, then it can be used advantageously in the optimization of (OF). This might also give rise to some multipliers that could be useful for statistical inference like $\boldsymbol{\theta}_\tau$, $\Psi_\tau(\boldsymbol{\theta}_\tau)$, $\mathbb{A}_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$, $\boldsymbol{s}_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$, and $c_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$, possibly considered as functions of $\tau$ and $g$. That is to say that the choice of $g$ matters in general and may have a huge impact on required moment assumptions as well as on the robustness and rigidity of the resulting elliptical regression quantile contours. In fact, the parametrization of quantile characteristics $\mathbb{A}_\tau$, $\boldsymbol{s}_\tau$, and $c_\tau$ is so important that it is repeatedly discussed throughout the next sections.

Unfortunately, the parametric elliptical regression $\tau$-quantiles are not uniquely defined in the instances when $\Psi_\tau(\boldsymbol{\theta})$ attains multiple global minima, which is typical of all nonlinear regression estimators; see [21] for a slightly more detailed discussion of that in the generic multivariate case.

If the lack of robustness is not an issue, then $g_I(t) = t$ seems the best choice because it can often be reasonably expected to minimize the number of local minima of (OF) as well as the overall computational burden. This choice also produces the very special uniquely defined elliptical regression quantiles described and illustrated in [17]. If robustness is of high priority, then one should choose either $g(t) = t^\alpha$ for $\alpha < 1$ to preserve affine equivariance or perhaps $g$ equal to a simple, bounded, and easy to compute function behaving like the identity function close to zero. However, if $\alpha < 0.5$ or $g$ is bounded, then the objective function (OF) may easily become misbehaving. This is why such choices cannot be recommended before such behavior and its consequences are fully clarified.

Obviously, the elliptical regression quantiles handle response outliers better than the design ones, because their robustness to design outliers may remain in question even for a bounded $g$ due to the possible negative impact of $c_\tau(\boldsymbol{\theta}, \boldsymbol{z})$. This defect is unpleasant although $c_\tau(\boldsymbol{\theta}, \boldsymbol{z})$ unbounded in $\boldsymbol{z}$ need not always spoil the robustness too much and although it can be bounded easily by means of a suitable parametrization; see Figure 3 for a result of such an attempt.

The definition of the parametric elliptical regression quantiles is so general that one can hardly say anything special about them without further assumptions. The next section attempts to point out some of their favorable properties without sacrificing too much generality. The following terminology then comes in handy.

**Definition 2.2.** The parametrization of the elliptical regression $\tau$-$g$-quantiles is called:

- **separable** if $\boldsymbol{\theta} = (\boldsymbol{\theta}'_{\boldsymbol{s}}, \boldsymbol{\theta}'_{\mathbb{A}}, \boldsymbol{\theta}'_c)'$ and $\boldsymbol{s}_\tau(\boldsymbol{\theta})$, $\mathbb{A}_\tau(\boldsymbol{\theta})$, and $c_\tau(\boldsymbol{\theta})$ really depend solely on $\boldsymbol{\theta}_{\boldsymbol{s}}$, $\boldsymbol{\theta}_{\mathbb{A}}$, and $\boldsymbol{\theta}_c$, respectively;
- **reducible** in $\boldsymbol{s}_\tau$ if $\boldsymbol{s}_\tau(\boldsymbol{\theta}, \boldsymbol{z}) = \boldsymbol{s}^0_\tau + \boldsymbol{s}^1_\tau(\boldsymbol{\theta}, \boldsymbol{z})$ where $\boldsymbol{s}^1_\tau$ is some function, and $\boldsymbol{s}^0_\tau$ is an $m$-dimensional subvector of $\boldsymbol{\theta}$ in which $\mathbb{A}_\tau(\boldsymbol{\theta})$, $c_\tau(\boldsymbol{\theta})$, and $\boldsymbol{s}^1_\tau(\boldsymbol{\theta})$ are constant;
- **reducible** in $c_\tau$ if $c_\tau(\boldsymbol{\theta}, \boldsymbol{z}) = c^0_\tau + c^1_\tau(\boldsymbol{\theta}, \boldsymbol{z})$ where $c^1_\tau$ is some function, and $c^0_\tau$ is a scalar subvector of $\boldsymbol{\theta}$ in which $\boldsymbol{s}_\tau(\boldsymbol{\theta})$, $\mathbb{A}_\tau(\boldsymbol{\theta})$, and $c^1_\tau(\boldsymbol{\theta})$ are constant;
- **admissible** if there exists $\boldsymbol{\theta}^0_\tau \in \Theta_\tau$ such that

  $$\boldsymbol{s}_\tau(\boldsymbol{\theta}^0_\tau, \boldsymbol{z}) = \boldsymbol{s}^0_\tau(\boldsymbol{z}),\ \mathbb{A}_\tau(\boldsymbol{\theta}^0_\tau, \boldsymbol{z}) = \mathbb{A}^0_\tau(\boldsymbol{z}),\ \text{and } c_\tau(\boldsymbol{\theta}^0_\tau, \boldsymbol{z}) = c^0_\tau(\boldsymbol{z})$$

  for almost all $\boldsymbol{z}$ where $\boldsymbol{s}^0_\tau(\boldsymbol{z})$, $\mathbb{A}^0_\tau(\boldsymbol{z})$, and $c^0_\tau(\boldsymbol{z})$ describe a multivariate elliptical $\tau$-$g$-quantile of the conditional distribution of $\boldsymbol{Y}$ given $\boldsymbol{Z} = \boldsymbol{z}$, as defined in [21]. It means that $\boldsymbol{s}^0_\tau(\boldsymbol{z})$, $\mathbb{A}^0_\tau(\boldsymbol{z})$, and $c^0_\tau(\boldsymbol{z})$ jointly minimize the expectation (with respect to the conditional distribution)

  $$\mathrm{E}_{\boldsymbol{Y}|\boldsymbol{Z}=\boldsymbol{z}}\, \rho_\tau\Big(g\big((\boldsymbol{Y} - \boldsymbol{s})'\mathbb{A}(\boldsymbol{Y} - \boldsymbol{s})\big) - c\Big)$$

  subject to the constraints that $\mathbb{A}$ is positive semidefinite and $\det(\mathbb{A}) = 1$.

The parametrization is therefore admissible if there exists $\boldsymbol{\theta}^0_\tau \in \Theta_\tau$ such that the $\boldsymbol{z}$-cuts of the corresponding elliptical regression $\tau$-$g$-quantile are equal to multivariate $\tau$-$g$-quantiles of the conditional distributions of $\boldsymbol{Y}$ given $\boldsymbol{Z} = \boldsymbol{z}$ for almost all $\boldsymbol{z}$.

**Example 2.1.** Consider $\tau \in (0, 1)$ and $(\boldsymbol{Y}', \boldsymbol{Z}')'$ with a multivariate normal distribution or with a multivariate elliptical distribution having all required moments finite. Then any separable parametrization of elliptical regression $\tau$-$g$-quantiles such that

1. $\mathbb{A}_\tau(\boldsymbol{\theta}, \boldsymbol{z})$, $\boldsymbol{\theta} \in \Theta_\tau$, does not depend on $\boldsymbol{z}$ and may become any positive definite matrix with unit determinant,
2. $\boldsymbol{s}_\tau(\boldsymbol{\theta}, \boldsymbol{z})$, $\boldsymbol{\theta} \in \Theta_\tau$, includes any affine function of $\boldsymbol{z}$, and
3. $c_\tau(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta_\tau$, does not depend on $\boldsymbol{z}$ and may attain any positive value,

is admissible for any permitted $g$ if it leads to the uniquely defined elliptical regression $\tau$-$g$-quantile; see [13] and Theorem 3.5 below.

## 3. BASIC PROPERTIES

The justification for elliptical regression quantiles is based on their good properties in the special location case, resulting from the necessary gradient conditions of [21]. The conditions play such a prominent role that they deserve to be paraphrased below using current terminology:

**Theorem 3.1.** *Consider the special location case (without regressors) when* $(\boldsymbol{Y}', \boldsymbol{Z}') = \boldsymbol{Y}'$ *and the parameters* $\boldsymbol{s}_\tau$, $\mathbb{A}_\tau$, *and* $c_\tau$ *are constant. Then the elliptical* $\tau$-$g$*-quantiles must satisfy the necessary conditions (1) to (4) of [21] that translate to*

$$1 = \det(\mathbb{A}_\tau)\,, \tag{3.1}$$

$$0 = P\big((\boldsymbol{Y}', \boldsymbol{Z}')' \in \mathcal{E}^-_{g,\tau}\big) - \tau\,, \tag{3.2}$$

$$\mathbf{0} = \frac{1}{1-\tau}\,\mathrm{E}\Big[\gamma\,\boldsymbol{R}_\tau\,\mathrm{I}_{[(\boldsymbol{Y}',\boldsymbol{Z}')'\in\mathcal{E}^+_{g,\tau}]}\Big] - \frac{1}{\tau}\,\mathrm{E}\Big[\gamma\,\boldsymbol{R}_\tau\,\mathrm{I}_{[(\boldsymbol{Y}',\boldsymbol{Z}')'\in\mathcal{E}^-_{g,\tau}]}\Big]\,, \tag{3.3}$$

*and*

$$L_\tau \frac{\det(\mathbb{A}_\tau)}{\tau(1-\tau)}\,\mathbb{A}_\tau^{-1} = \frac{1}{1-\tau}\,\mathrm{E}\Big[\gamma\,\boldsymbol{R}_\tau\boldsymbol{R}_\tau'\,\mathrm{I}_{[(\boldsymbol{Y}',\boldsymbol{Z}')'\in\mathcal{E}^+_{g,\tau}]}\Big] - \frac{1}{\tau}\,\mathrm{E}\Big[\gamma\,\boldsymbol{R}_\tau\boldsymbol{R}_\tau'\,\mathrm{I}_{[(\boldsymbol{Y}',\boldsymbol{Z}')'\in\mathcal{E}^-_{g,\tau}]}\Big]\,, \tag{3.4}$$

*where* $\mathbb{A}_\tau$ *is assumed symmetric positive semidefinite,* $L_\tau$ *is the Lagrange multiplier corresponding to the constraint* $-\det(\mathbb{A}_\tau) + 1 = 0$, $\boldsymbol{R}_\tau = \boldsymbol{Y} - \boldsymbol{s}_\tau$, $\dot{g}(t) := \partial g(t)/\partial t$, *and* $\gamma = \dot{g}(\boldsymbol{R}_\tau'\mathbb{A}_\tau\boldsymbol{R}_\tau)$.

The probability interpretation of the location elliptical quantiles then results from (3.2). If $g = g_I$, then $\gamma = 1$ and the conditions simplify considerably and become easy to interpret; see [21] for further details.

In the general regression context considered here, $\boldsymbol{s}_\tau$, $\mathbb{A}_\tau$, and $c_\tau$ may depend on $\boldsymbol{z}$ and on the common underlying parameter $\boldsymbol{\theta}$. Consequently, one should derive (OF) as a compound function and the derivatives of $\boldsymbol{s}_\tau$, $\mathbb{A}_\tau$, and $c_\tau$ with respect to $\boldsymbol{\theta}$ should also enter the scene.

If the properties of elliptical regression quantiles should naturally generalize those of the location ones, then only separable parametrizations reducible both in $c_\tau$ and $\boldsymbol{s}_\tau$ should be considered.

The next theorem summarizes some obvious special cases.

**Theorem 3.2.** *If the parametrization of the elliptical regression* $\tau$*-quantiles*

- *is reducible in* $c_\tau$, *then (3.2) holds;*
- *is reducible in* $\boldsymbol{s}_\tau$ *with* $\boldsymbol{z}$*-independent* $\mathbb{A}_\tau$, *then (3.3) holds;*
- *is separable and* $c_\tau = \boldsymbol{\theta}_L'\boldsymbol{z} + c^I_\tau(\boldsymbol{\theta}_c, \boldsymbol{z})$ *where* $\boldsymbol{\theta}_L$ *is a subvector of* $\boldsymbol{\theta}_c$ *in which* $c^I_\tau$ *is constant, then*

$$\mathbf{0} = \frac{1}{1-\tau}\,\mathrm{E}\Big[\boldsymbol{Z}\,\mathrm{I}_{[(\boldsymbol{Y}',\boldsymbol{Z}')'\in\mathcal{E}^+_{g,\tau}]}\Big] - \frac{1}{\tau}\,\mathrm{E}\Big[\boldsymbol{Z}\,\mathrm{I}_{[(\boldsymbol{Y}',\boldsymbol{Z}')'\in\mathcal{E}^-_{g,\tau}]}\Big]\,.$$

Assume that all the three conditions are satisfied. Then the population parametric elliptical regression quantiles have a clear probability interpretation, $\mathcal{E}^-_{g,\tau}$ is nonempty for $\tau > 0$, and the centers of probability mass of $\mathcal{E}^-_{g,\tau}(\boldsymbol{Y}, \boldsymbol{Z})$ and $\mathcal{E}^+_{g,\tau}(\boldsymbol{Y}, \boldsymbol{Z})$ have the same $\boldsymbol{z}$-coordinates. The second claim then meaningfully links the probability mass centers of scaled residuals $\gamma(\boldsymbol{Y} - \boldsymbol{s}_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{Z}))$ corresponding to the regression observations in $\mathcal{E}^-_{g,\tau}(\boldsymbol{Y}, \boldsymbol{Z})$ and $\mathcal{E}^+_{g,\tau}(\boldsymbol{Y}, \boldsymbol{Z})$.

Every reasonable multivariate quantile regression concept should also exhibit good equivariance properties. The parametric elliptical quantile regression need not be an exception in this regard. What really matters is how $\boldsymbol{s}_\tau(\boldsymbol{\theta}_\tau)$, $\mathbb{A}_\tau(\boldsymbol{\theta}_\tau)$, and $c_\tau(\boldsymbol{\theta}_\tau)$ change with the transformations of $\boldsymbol{Y}$, and this follows directly from the location case of [21].

**Definition 3.1.** The parametrization of elliptical regression $\tau$-$g$-quantiles is called affine equivariant if $g(t) = t^r$ for some $r > 0$ and if, for any $\boldsymbol{a} \in \mathbb{R}^m$, any regular $m \times m$ matrix $\mathbb{B}$ (with determinant $d$), and any $\boldsymbol{\theta} \in \Theta_\tau$, there exists $\boldsymbol{\theta}_{\mathbb{B},\boldsymbol{a},d} \in \Theta_\tau$ such that

$$\mathbb{A}_\tau(\boldsymbol{\theta}_{\mathbb{B},\boldsymbol{a},d}, \boldsymbol{z}) = d^2 \big(\mathbb{B}^{-1}\big)' \mathbb{A}_\tau(\boldsymbol{\theta}, \boldsymbol{z})\, \mathbb{B}^{-1}, \tag{3.5}$$

$$\boldsymbol{s}_\tau(\boldsymbol{\theta}_{\mathbb{B},\boldsymbol{a},d}, \boldsymbol{z}) = \boldsymbol{a} + \mathbb{B}\, \boldsymbol{s}_\tau(\boldsymbol{\theta}, \boldsymbol{z}), \tag{3.6}$$

and

$$c_\tau(\boldsymbol{\theta}_{\mathbb{B},\boldsymbol{a},d}, \boldsymbol{z}) = g\Big(d^2 g^{-1}\big(c_\tau(\boldsymbol{\theta}, \boldsymbol{z})\big)\Big) \tag{3.7}$$

for all $\boldsymbol{z}$. If (3.5), (3.6) and (3.7) hold for $d = 1$, then the parametrization is called shift and rotation equivariant, even if $g$ is not a polynomial.

**Theorem 3.3.** *If the parametrization of elliptical regression $\tau$-quantiles is affine equivariant, then the resulting elliptical regression $\tau$-quantiles are affine equivariant. If it is shift and rotation equivariant, then the resulting elliptical regression $\tau$-quantiles are shift and rotation equivariant.*

**Proof:** If $\boldsymbol{\theta} \in \Theta_\tau$ minimizes (OF) for random vector $(\boldsymbol{Y}', \boldsymbol{Z}')' \in \mathbb{R}^{m+p}$, then corresponding $\boldsymbol{\theta}_{\mathbb{B},\boldsymbol{a},d} \in \Theta_\tau$ from the above definition of the equivariant parametrization obviously minimizes (OF) for random vector $\big((\boldsymbol{a} + \mathbb{B}\boldsymbol{Y})', \boldsymbol{Z}'\big)' \in \mathbb{R}^{m+p}$ for any $\boldsymbol{a} \in \mathbb{R}^m$ and any regular $m \times m$ matrix $\mathbb{B}$ with determinant $d$. $\square$

In other words, if the elliptical regression $\tau$-quantile of $(\boldsymbol{Y}', \boldsymbol{Z}')'$ is parametrized with $\mathbb{A}_\tau$, $\boldsymbol{s}_\tau$, and $c_\tau$ by means of an affine equivariant parametrization, then the elliptical regression $\tau$-quantile of $\big((\boldsymbol{a} + \mathbb{B}\boldsymbol{Y})', \boldsymbol{Z}'\big)'$ can be parametrized with $d^2\big(\mathbb{B}^{-1}\big)'\mathbb{A}_\tau\mathbb{B}^{-1}$, $\boldsymbol{a} + \mathbb{B}\boldsymbol{s}_\tau$, and $g\big(d^2 g^{-1}\big(c_\tau(\boldsymbol{\theta}, \boldsymbol{z})\big)\big)$.

The graph of $\Psi_\tau(\boldsymbol{\theta})$ crucially influences the process of optimization. The following consequences of convex calculus might serve as a guidance for choosing $g$ and minimizing the troubles with the optimization of $\Psi_\tau(\boldsymbol{\theta})$.

**Theorem 3.4.** *Assume a separable parametrization of the elliptical regression $\tau$-$g$-quantiles with $\boldsymbol{\theta} = (\boldsymbol{\theta}'_{\boldsymbol{s}}, \boldsymbol{\theta}'_{\mathbb{A}}, \boldsymbol{\theta}'_c)'$.*

- *If $g = g_I$, then $\Psi_\tau$ is convex in $\mathbb{A}_\tau$.*
- *If $c_\tau$ is linear in $\boldsymbol{\theta}_c$, then $\Psi_\tau(\boldsymbol{\theta})$ is convex in $\boldsymbol{\theta}_c$.*

In fact, $g = g_I$ may easily lead to uniquely defined parametric elliptical regression quantiles; see [17].

Generally speaking, the good properties of multivariate elliptical quantiles extend to the elliptical regression quantiles with admissible and affine equivariant parametrizations.

**Theorem 3.5.** *Let $\tau \in (0,1)$ and $f(\boldsymbol{y},\boldsymbol{z}) = f_1(\boldsymbol{y}|\boldsymbol{z})\, f_2(\boldsymbol{z})$ be the density of $(\boldsymbol{Y}',\boldsymbol{Z}')' \in \mathbb{R}^{m+p}$ where $f_2(\boldsymbol{z})$ is the marginal density of $\boldsymbol{Z}$ and $f_1(\boldsymbol{y}|\boldsymbol{z})$ is the regularized version of the density of the conditional distribution of $\boldsymbol{Y}$ given $\boldsymbol{Z} = \boldsymbol{z}$ that is assumed to exist.*

*If the parametrization $\mathbb{A}_\tau(\boldsymbol{\theta},\boldsymbol{z})$, $\boldsymbol{s}_\tau(\boldsymbol{\theta},\boldsymbol{z})$, and $c_\tau(\boldsymbol{\theta},\boldsymbol{z})$ of the elliptical regression $\tau$-quantile is admissible, then there exists $\boldsymbol{\theta}_\tau \in \Theta_\tau$ minimizing (OF). If for any orthonormal matrix $\mathbb{O}$ there exists $\widetilde{\boldsymbol{\theta}}_\tau(\mathbb{O}) \in \Theta_\tau$ such that $\mathbb{A}_\tau(\widetilde{\boldsymbol{\theta}}_\tau(\mathbb{O}),\boldsymbol{z}) = \mathbb{O}'\mathbb{A}_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z})\,\mathbb{O}$, $c_\tau(\widetilde{\boldsymbol{\theta}}_\tau(\mathbb{O}),\boldsymbol{z}) = c_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z})$, and $\boldsymbol{s}_\tau(\widetilde{\boldsymbol{\theta}}_\tau(\mathbb{O}),\boldsymbol{z}) = \boldsymbol{\mu}(\boldsymbol{z}) + \mathbb{O}'\big(\boldsymbol{s}_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z}) - \boldsymbol{\mu}(\boldsymbol{z})\big)$ for the particular $\boldsymbol{\mu}$ appearing below, and*

[**1**] *if $f_1(\boldsymbol{y}|\boldsymbol{z}) = f_1\big(\boldsymbol{\mu}(\boldsymbol{z}) + \mathbb{O}(\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{z}))\,|\,\boldsymbol{z}\big)$ for some function $\boldsymbol{\mu} = (\mu_1,\dots,\mu_m)'$ and for an orthonormal matrix $\mathbb{O} = \mathbb{O}^{-1'}$, then there exists an elliptical regression $\tau$-quantile parametrized with $\mathbb{A}_\tau(\widetilde{\boldsymbol{\theta}}_\tau(\mathbb{O}),\boldsymbol{z})$, $\boldsymbol{s}_\tau(\widetilde{\boldsymbol{\theta}}_\tau(\mathbb{O}),\boldsymbol{z})$, and $c_\tau(\widetilde{\boldsymbol{\theta}}_\tau(\mathbb{O}),\boldsymbol{z})$.*

*If the elliptical regression $\tau$-quantile is moreover uniquely defined, then*

[**2**] *if $\boldsymbol{s}_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z}) = (s_1,\dots,s_m)(\boldsymbol{z})'$, $\mathbb{A}_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z}) = (a_{ij}(\boldsymbol{z}))_{i,j=1}^m$, and $f_1(\boldsymbol{y}|\boldsymbol{z}) = f_1\big(\boldsymbol{\mu}(\boldsymbol{z}) + \mathbb{J}(\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{z}))\,|\,\boldsymbol{z}\big)$ for all $\boldsymbol{z}$ and a sign-change matrix $\mathbb{J} = \mathbb{J}' = \mathbb{J}^{-1} = \mathrm{diag}(j_1,\dots,j_m)$ with diagonal elements $\pm 1$, then $s_i(\boldsymbol{z}) = \mu_i(\boldsymbol{z})$ whenever $j_i = -1$, $i \in \{1,\dots,m\}$, and $a_{ij}(\boldsymbol{z}) = 0$ whenever $j_i j_j = -1$, $i,j \in \{1,\dots,m\}$;*

[**3**] *if all the conditional distributions of $\boldsymbol{Y}$ given $\boldsymbol{Z} = \boldsymbol{z}$ are centrally symmetric around their center of symmetry $\boldsymbol{\mu}(\boldsymbol{z})$, then $\boldsymbol{s}_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z}) = \boldsymbol{\mu}(\boldsymbol{z})$;*

[**4**] *if all the conditional distributions of $\boldsymbol{Y}$ given $\boldsymbol{Z} = \boldsymbol{z}$ centered with $\boldsymbol{\mu}(\boldsymbol{z})$ are symmetric around a common hyperplane $H$, then $\boldsymbol{s}_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z}) - \boldsymbol{\mu}(\boldsymbol{z})$ lies on $H$;*

[**5**] *if all the conditional distributions of $\boldsymbol{Y}$ given $\boldsymbol{Z} = \boldsymbol{z}$ centered with $\boldsymbol{\mu}(\boldsymbol{z})$ are symmetric along a common axis $o$, then $\boldsymbol{s}_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z}) - \boldsymbol{\mu}(\boldsymbol{z})$ lies on that axis.*

**Proof:** As for [1], the assumed admissible parametrization guarantees that there exists $\boldsymbol{\theta}_\tau \in \Theta_\tau$ such that $\mathbb{A}_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z})$, $\boldsymbol{s}_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z})$, and $c_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z})$ minimize $\Phi_\tau^{\boldsymbol{z}}(\mathbb{A},\boldsymbol{s},c) := \mathrm{E}_{\boldsymbol{Y}|\boldsymbol{Z}=\boldsymbol{z}}\, \rho_\tau\big(g\big((\boldsymbol{Y}-\boldsymbol{s})'\,\mathbb{A}(\boldsymbol{Y}-\boldsymbol{s})\big) - c\big)$ for almost all $\boldsymbol{z}$. Therefore, they minimize (OF) as well. The assumption on the conditional density further implies $\Phi_\tau^{\boldsymbol{z}}(\mathbb{A}_\tau,\boldsymbol{s}_\tau,c_\tau) = \Phi_\tau^{\boldsymbol{z}}\big(\mathbb{O}'\mathbb{A}_\tau\mathbb{O}, \boldsymbol{\mu}(\boldsymbol{z}) + \mathbb{O}'(\boldsymbol{s}_\tau - \boldsymbol{\mu}(\boldsymbol{z})), c_\tau\big)$, and thus $\mathbb{O}'\mathbb{A}_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z})\,\mathbb{O}$, $\boldsymbol{\mu}(\boldsymbol{z}) + \mathbb{O}'\big(\boldsymbol{s}_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z}) - \boldsymbol{\mu}(\boldsymbol{z})\big)$ and $c_\tau(\boldsymbol{\theta}_\tau,\boldsymbol{z})$ also minimize not only the same conditional expectation for almost all $\boldsymbol{z}$, but also (OF) as well, and, therefore, they also describe an elliptical regression $\tau$-quantile thanks to the assumed existence of $\widetilde{\boldsymbol{\theta}}_\tau(\mathbb{O})$.

As for [2], it follows directly from [1] because matrix $\mathbb{J}$ is orthonormal. Only the two elliptical regression $\tau$-quantiles from [1] must now coincide due to the uniqueness assumption. This fact implies $s_i(\boldsymbol{z}) = \mu_i(\boldsymbol{z})$ whenever $j_i = -1$, and $a_{ij}(\boldsymbol{z}) = 0$ whenever $j_i j_j = -1$, $i,j \in \{1,\dots,m\}$. Furthermore, [2] implies [3] for $\mathbb{J} = -\mathbb{I}$. The rest ([4] and [5]) analogously results from [1] and [2] for certain orthonormal matrices. □

**Note 3.1.** In [1], [2], and [3], it would be enough to assume the existence of $\widetilde{\boldsymbol{\theta}}_\tau(\mathbb{O}) \in \Theta_\tau$ only for the particular orthonormal matrices $\mathbb{O}$ considered there. In fact, the statements [2] to [5] could be proved directly by generalizing the location case with similar behavior regarding symmetry, only with the requirement of an admissible parametrization and without any need of $\widetilde{\boldsymbol{\theta}}_\tau(\mathbb{O})$ for some orthonormal matrices $\mathbb{O}$.

**Note 3.2.** The somewhat analogous Theorem 1 of [21] and its proof unfortunately contain a couple of misprints and one error. First, any occurrence of $\mathbb{O}\boldsymbol{s}_\tau$ should be replaced with $\mathbb{O}'\boldsymbol{s}_\tau$ there. Second, the proof should apply (2) to (6), not (2)–(6). And most importantly, the natural behavior of generalized elliptical quantiles under affine transformations of the response vector, postulated by Theorem 1 (1), is there falsely interpreted as full affine equivariance for any function $g$, which invalidates the proofs of further statements (3), (4), (5), and (10). While the generalized elliptical quantiles are always shift and rotation equivariant, they are certain to be fully affine equivariant only for $g(t) = t^\alpha$, $\alpha > 0$. Consequently, the statements (3), (4), (5), and (10) there hold only for such functions $g$ or for spherical distributions. The claims (6)–(9) there really require only rotation and shift equivariance and, therefore, remain valid for any function $g$ as they stand.

The uniqueness assumption used in Theorem 3.5 is not as severe as it might seem at first sight. That is to say that what really matters is only the uniqueness of $\mathbb{A}_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$, $\boldsymbol{s}_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$ and $\boldsymbol{c}_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$ in the population case.

Any admissible parametrization by definition guarantees the existence of such $\boldsymbol{\theta}_0 \in \Theta_\tau$ that (for almost all $\boldsymbol{z}$) minimizes the (non-negative finite) conditional expectation of $\rho_\tau(h_\tau(\boldsymbol{\theta}, \boldsymbol{Y}, \boldsymbol{Z}))$ (with respect to the conditional distribution of $\boldsymbol{Y}$ given $\boldsymbol{Z} = \boldsymbol{z}$). This implies that the same $\boldsymbol{\theta}_0$ also minimizes its unconditional (finite) expectation (OF). Therefore, the parameter vector $\boldsymbol{\theta}_0 \in \Theta_\tau$ also defines an elliptical regression $\tau$-quantile that is uniquely defined if all the purely multivariate elliptical $\tau$-quantiles of $\mathcal{L}(\boldsymbol{Y} \,|\, \boldsymbol{Z} = \boldsymbol{z})$ are uniquely defined. The uniqueness of multivariate elliptical $\tau$-quantiles has been studied in [20, 21] and established for $g(t) = t$ under very mild conditions. Consequently, the aforementioned considerations extend the uniqueness result even to elliptical regression quantiles with $g(t) = t$ and admissible parametrizations. This is why $g(t) = t$ is generally preferred to other possibilities for the time being.

Unfortunately, ill-specified models for elliptical regression quantiles generally need not lead to a unique solution even for $g(t) = t$. This is typical of all nonlinear regression methods. Nevertheless, there exist certain natural parametrizations with $g(t) = t$ that lead to unique elliptical regression quantiles even if the model is misspecified; see [17].

## 4. THE ART OF PARAMETRIZATION

The parametrization of $\boldsymbol{s}_\tau$ follows directly from available preconceptions regarding the multivariate trend, and that of $c_\tau$ also often results from the context quite easily. One choice can be nevertheless much better than its formal equivalents from the computational point of view; see Section 6.

On the contrary, it need not be that clear how to parametrize $\mathbb{A}_\tau$ to keep it positive definite with unit determinant so that one could avoid all the restrictions and constrained optimization. In the case of bivariate responses with $m = 2$, there are several possibilities at hand, e.g.

$$\mathbb{A}_\tau(\boldsymbol{\theta}, \boldsymbol{z}) = \begin{pmatrix} a_{11}^2 & a_{12} \\ a_{12} & (1 + a_{12}^2)/a_{11}^2 \end{pmatrix}, \tag{4.1}$$

$$\mathbb{A}_\tau(\boldsymbol{\theta}, \boldsymbol{z}) = \begin{pmatrix} c_1 & c_2 \\ 0 & \frac{1}{c_1} \end{pmatrix}' \begin{pmatrix} c_1 & c_2 \\ 0 & \frac{1}{c_1} \end{pmatrix} = \begin{pmatrix} c_1^2 & c_1 c_2 \\ c_1 c_2 & c_2^2 + \frac{1}{c_1^2} \end{pmatrix}, \tag{4.2}$$

or

$$\mathbb{A}_\tau(\boldsymbol{\theta}, \boldsymbol{z}) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}' \begin{pmatrix} d^2 & 0 \\ 0 & \frac{1}{d^2} \end{pmatrix} \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}, \tag{4.3}$$

where the obvious dependence of $a_{11}$, $a_{12}$, $c_1$, $c_2$, $\alpha$, and $d^2$ on $\tau$, $\boldsymbol{\theta}$, and $\boldsymbol{z}$ is not emphasized for the sake of brevity. Of course, one could also consider $\exp(a_{11})$ and $\exp(d)$ instead of $a_{11}^2$ and $d^2$, not to mention other alternatives in the same spirit.

Clearly, (4.1) is the most straightforward possibility but it can hardly be generalized beyond dimension $m = 2$ or $m = 3$. On the other hand, (4.2) follows from the Choleski decomposition advocated in [21] and it can be easily adjusted to any dimension of the responses. The third example (4.3) results from the spectral decomposition and it also can be extended to general multivariate response settings, though in a rather complicated way.

The optimal choice of parametrization for $\mathbb{A}_\tau$ crucially depends on the type of expected heteroscedasticity. The spectral decomposition in (4.3) appears very appealing due to its easy and natural interpretation. Unfortunately, such a parametrization of a positive definite matrix is not unique without further assumptions regarding the angles and/or the diagonal elements of the sandwiched matrix. Sometimes one can give up the uniqueness, find a solution, and then transform it to a canonical form without any harm. One could also use the well-worn tricks how to enforce one parameter higher than the other or in a certain range. The choices may depend on the expected model, which shifts the modeling from a boring routine to sophisticated art.

In the cases of homoscedasticity and multiplicative heteroscedasticity described below and corresponding to constant $\mathbb{A}_\tau$, one can simply avoid all such problems by using the parametrization based on the Choleski decomposition, which is generally recommended in such situations.

## 5. CLASSIFICATION OF HETEROSCEDASTICITY

Assume that a correctly specified elliptical quantile regression model for bivariate responses leads to a unique solution $\mathbb{A}_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$, $\boldsymbol{s}_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$, and $c_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$, with $\mathbb{A}_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$ parametrized by means of $\alpha_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$ and $d_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$ as in (4.3). Then it makes sense to speak of $\tau$-level homoscedasticity when $\alpha_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$, $c_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$, and $d_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$ are all independent of $\boldsymbol{z}$. Furthermore, it is possible to distinguish three canonical $\tau$-level heteroscedastic patterns corresponding to the cases when only one of the characteristics $\alpha_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$, $c_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$, and $d_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$ depends on $\boldsymbol{z}$:

(**1**) rotational heteroscedasticity (if only $\alpha_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$ is $\boldsymbol{z}$-dependent),

(**2**) multiplicative (or scale) heteroscedasticity (if only $c_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$ is $\boldsymbol{z}$-dependent), and

(**3**) proportional heteroscedasticity (if only $d_\tau(\boldsymbol{\theta}_\tau, \boldsymbol{z})$ is $\boldsymbol{z}$-dependent).

Any type of bivariate heteroscedasticity can then be decomposed into the three canonical forms. See Figure 1 for an illustration of this classification.

If these heteroscedastic patterns are observed for all $\tau \in (0,1)$, then one can speak of $\tau$-independent heteroscedastic patterns. If they are observed only locally in $\tau$ or $\boldsymbol{z}$, then one can speak of local heteroscedastic patterns. This terminology can be adopted even informally when the true underlying model is unknown but its heteroscedastic profile slightly resembles that of elliptical quantile regression.

The situation becomes more complicated in case of multivariate responses, but even then the classification can still be used for any couple of their coordinates and the terms like overall rotational/proportional/multiplicative heteroscedasticity still make perfect sense.

Although the multiplicative heteroscedasticity seems by far the most common, the others are not necessarily extinct but maybe only hidden because the ways available for their detection and modeling are rather limited and unpopular, at least for the time being. For example, the rotational heteroscedasticity may be dormant in the data observed by the satellites orbiting the Earth. And it is demonstrated below in Section 9 that it might be present even in biometric data.

## 6. COMPUTATION

The sample elliptical regression $\tau$-$g$-quantiles can be obtained directly from the definition if the expectation in (OF) is taken with respect to the discrete empirical probability distribution. Consider $n$ responses $\boldsymbol{Y}_i$'s accompanied with corresponding regressor vectors $\boldsymbol{Z}_i$'s, $i = 1, ..., n$, from the population distribution assumed above. Even if all the constraints on $\mathbb{A}_\tau$ are removed in the way described in Section 4, then it still remains to solve the unconstrained optimization problem

$$\min_{\boldsymbol{\theta}} \sum_{i=1}^{n} \rho_\tau\big(h_\tau(\boldsymbol{\theta}, \mathbf{Y}_i, \boldsymbol{Z}_i)\big)$$

for appropriate $h_\tau$ where the objective function is generally neither smooth nor convex. Of course, it could be done with a suitable general solver for non-convex optimization. Fortunately, this problem can also be viewed as a nonlinear quantile regression task with zero responses and regressors $(\boldsymbol{Y}_i', \boldsymbol{Z}_i')'$, $i = 1, ..., n$, that has already been studied successfully, see [22], and can be solved for differentiable $h_\tau$ with the special algorithm developed in [24] whose Matlab implementation in ipqr.m, available at `http://sites.stat.psu.edu/~dhunter/code/qrmatlab`, had been tuned up and used for the computation of all the sample parametric elliptical regression $g$-quantiles presented in the next sections. In other words, the parametric elliptical regression quantiles can be computed like their location predecessors of [21].

Unfortunately, the algorithm of [24] must be initialized with a preliminary estimate of $\boldsymbol{\theta}_\tau$.

This is a stage when any available information about the estimated vector parameter can be employed advantageously. Of course, one should experiment with several wise choices of initial parameters and then choose the solution according to the final parameter estimators and corresponding values of the minimized objective function. If some not-so-complicated regression models were considered, then one might also fit each response component by means of single-response quantile regression and use the resulting parameter estimates to initialize the algorithm. A few multivariate quantile cuts obtained from other multi-response quantile regression method(s) could also be mined for some information leading to the initial parameter estimates.

The parametrization of the problem also matters as one can lead to the successful end much more quickly and easily than another. From this point of view, it is strongly recommended to avoid nonlinearities whenever possible. If the Jacobian derived from $h_\tau$ is singular from the very beginning or becomes singular or close to singular during the computation, then insuperable numerical problems can be expected, which also speaks for using well-thought-out parametrizations and parameter initializations. For example, such a situation may happen for $d^2 = 1$ if the parametrization (4.3) is used for $\mathbb{A}_\tau$.

The computational side of many nonlinear regression methods is not ideal and the parametric elliptical quantile regression is no exception in this regard. But one can hardly hope for anything else if the model is genuinely nonlinear and non-convex in its parameters.

## 7. MODEL VALIDATION

This section suggests a few heuristic ways how to validate the resulting elliptical quantile regression models before the topic is treated elsewhere in full detail and exactness. The first two are commonly used in the ordinary least squares regression.

Suppose that $n$ regression observations $(\boldsymbol{Y}_i', \boldsymbol{Z}_i')'$, $i = 1, ..., n$, were fitted with a generalized parametric elliptical ($\tau$-$g$-)quantile regression model leading to unique quantile parameter estimates $\mathbb{A}(\widehat{\boldsymbol{\theta}}, \boldsymbol{z})$, $\boldsymbol{s}(\widehat{\boldsymbol{\theta}}, \boldsymbol{z})$, $c(\widehat{\boldsymbol{\theta}}, \boldsymbol{z})$, and to homogenized (pseudo)residuals $r_i(\widehat{\boldsymbol{\theta}}) := h(\widehat{\boldsymbol{\theta}}, \boldsymbol{Y}_i, \boldsymbol{Z}_i)$, $i = 1, ..., n$; see Definition 2.1 for the origin of $h$.

One can then use the cross-validation approach to look for outliers or influential observations. In other words, the impact of some observation(s) can be evaluated by means of the differences $\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_-$, $\Psi(\widehat{\boldsymbol{\theta}}) - \Psi(\widehat{\boldsymbol{\theta}}_-)$, $c(\widehat{\boldsymbol{\theta}}, \boldsymbol{z}) - c(\widehat{\boldsymbol{\theta}}_-, \boldsymbol{z})$, $g^{-1}(c(\widehat{\boldsymbol{\theta}}, \boldsymbol{z})) - g^{-1}(c(\widehat{\boldsymbol{\theta}}_-, \boldsymbol{z}))$, $\boldsymbol{s}(\widehat{\boldsymbol{\theta}}, \boldsymbol{z}) - \boldsymbol{s}(\widehat{\boldsymbol{\theta}}_-, \boldsymbol{z})$, $\mathbb{A}(\widehat{\boldsymbol{\theta}}, \boldsymbol{z}) - \mathbb{A}(\widehat{\boldsymbol{\theta}}_-, \boldsymbol{z})$, $\mathbb{A}^{-1}(\widehat{\boldsymbol{\theta}}, \boldsymbol{z}) - \mathbb{A}^{-1}(\widehat{\boldsymbol{\theta}}_-, \boldsymbol{z})$, $r_i(\widehat{\boldsymbol{\theta}}) - r_i(\widehat{\boldsymbol{\theta}}_-)$, and their parts or norms where $\widehat{\boldsymbol{\theta}}_-$ is the quantile coefficient estimate obtained by excluding the suspected observation(s) from the sample. Of course, the differences of the whole quantile cuts corresponding to $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\theta}}_-$ could also be investigated. And it would be wise to consult such differences even in testing various submodels where the role of $\widehat{\boldsymbol{\theta}}_-$ would be played by the optimal estimate of $\boldsymbol{\theta}$ in the restricted model.

One could also inspect various charts to check the behavior of the homogenized (pseudo) residuals. In a well-specified model, they should be (roughly) mutually independent, identically distributed, and independent of the covariates (and also of the responses if all the conditional distributions were elliptical). For example, one may plot $r_i$ or $r_i^2$ on their lagged values and (the norms or components of) $\boldsymbol{Y}_i$ and $\boldsymbol{Z}_i$, $i = 1, ..., n$.

One could verify as well whether the estimated quantile cuts share their centers, axes, and hyperplanes of symmetry with the expected conditional distributions. The opposite might imply that the model assumptions were wrong, owing to Theorem 3.5.

If $c(\widehat{\boldsymbol{\theta}}, \boldsymbol{z})$ is unexpectedly negative for common regressor values, then there must be something wrong with the model specification too.

Finally, one might also validate the model by comparing the resulting quantile cuts with those obtained with another multivariate quantile regression method that requires even weaker assumptions and is still applicable to the data. Depending on the context, the benchmark or the nonparametric proposals of [26], [20], [14] or [4] could often serve the purpose quite well.

## 8. ILLUSTRATIONS

This section presents some pictures to support the claim that the parametric elliptical regression $g$-quantiles are indeed promising candidates for wide dissemination thanks to their many good properties. For the sake of simplicity, only the most often recommended natural choice $g_I(t) = t$ is considered hereinafter.

Unfortunately, the precise rules for choosing $g$ in different situations are still to be developed. For the time being, it only seems wise to scale the data properly before their analysis and then to use $g_I$ in the absence of outliers. The choice is also preferable from the computational point of view.

The examples below testify that the elliptical quantile regression can work well both for elliptical and non-elliptical underlying error distributions, and also for the number of observations $n$ as low as 99 and as high as 99 999. For the sake of simplicity and ease of presentation, the colors of both data points and quantile cuts are changing in dependence of the corresponding regressor values, and only bivariate responses with scalar regressors are considered. Nevertheless, there is no intrinsic restriction on the dimension of responses or regressors involved in the empirical model provided that the number of free model parameters is low relative to the total number of observations and not too large for the computation to terminate successfully.

The elliptical regression $\tau$-$g$-quantiles are parametrized by means of $\boldsymbol{s}_\tau$, $\mathbb{A}_\tau$, and $c_\tau$. In the examples, $\mathbb{A}_\tau$ is always considered in its spectral decomposition (4.3) described by $d_\tau^2$ and $\alpha_\tau$, although less complicated parametrizations of $\mathbb{A}_\tau$ should be generally preferred for models with constant $\mathbb{A}_\tau$; see Section 4 for the discussion of some possibilities.

Figure 1 is included to demonstrate that parametric elliptical $g$-quantile regression is suitable for both small and large data sets and for capturing various kinds of heteroscedasticity.

Figure 2 illustrates another key advantage of elliptical regression $g$-quantiles, namely their ability to easily incorporate many types of a priori information regarding the model parameters. Last but not least, Figure 3 indicates that the concept of parametric elliptical regression quantiles is not bound to linear regression settings and can be used even for fitting highly complicated nonlinear models.

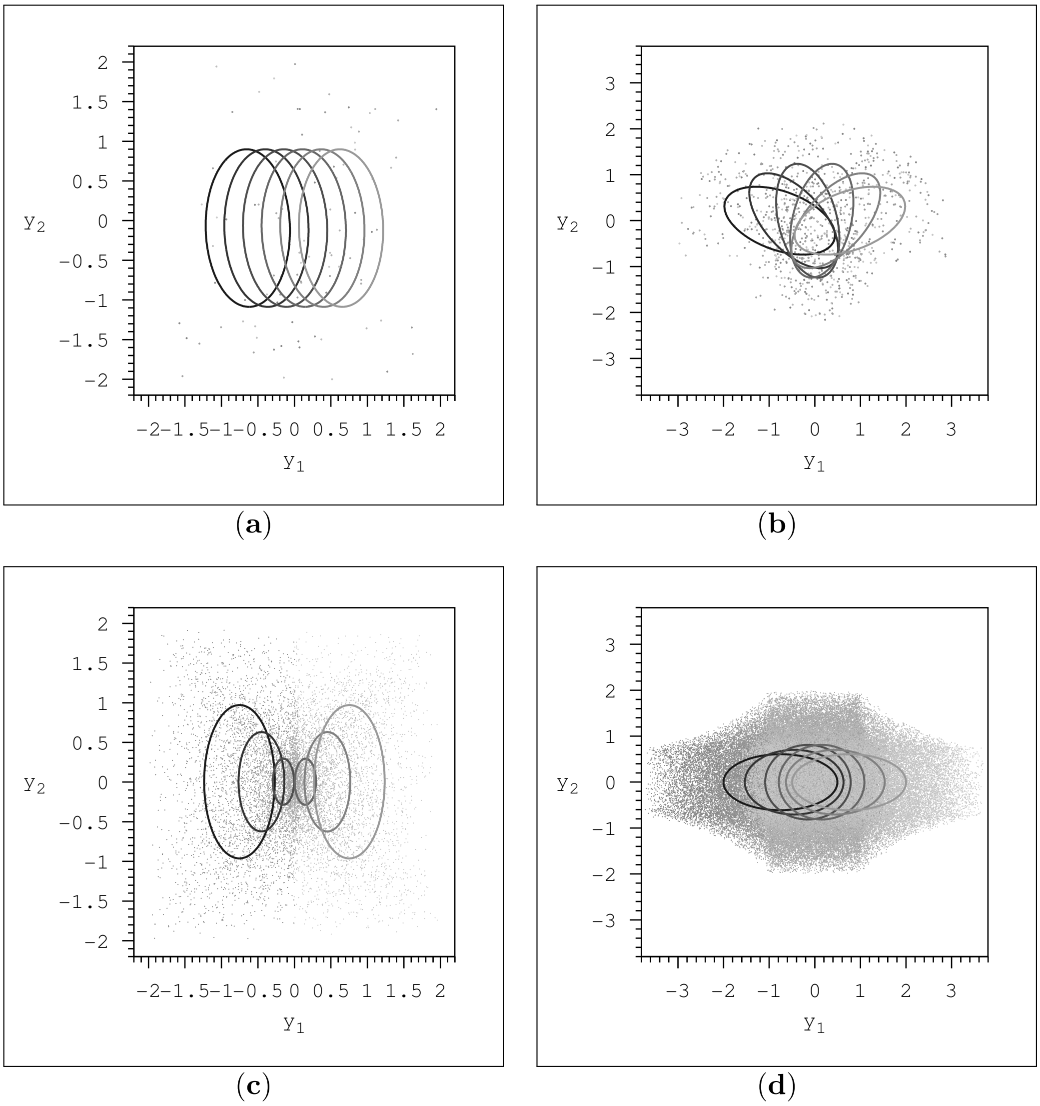



**Figure 1**: **Classification of heteroscedasticity in $\mathbb{R}^2$.** The plots illustrate four basic patterns of heteroscedasticity in $\mathbb{R}^2$ with elliptical regression 0.3-$g_I$-quantile cuts computed for six equidistant reference points $z_0 = -0.75, -0.45, ..., 0.75$ from $n$ regression observations $(Y_1, Y_2, Z)$ generated by the regression model $(Y_1, Y_2)' = (Z, 0)' + q(\boldsymbol{\varepsilon})$, $Z \sim U([-1,1])$ is independent of $\boldsymbol{\varepsilon} \sim U([-1,1]) \times U([-2,2])$:
(**a**) no heteroscedasticity $\big[n = 99,\ q(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}\big]$,
(**b**) rotational heteroscedasticity $\big[n = 999,\ q(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}'\mathbb{P}$ where $\operatorname{vec}(\mathbb{P})' = \big(\cos(\pi Z/2), \sin(\pi Z/2), -\sin(\pi Z/2), \cos(\pi Z/2)\big)\big]$,
(**c**) multiplicative heteroscedasticity $\big[n = 9\,999,\ q(\boldsymbol{\varepsilon}) = (0.1 + 0.9|Z|)\,\boldsymbol{\varepsilon}\big]$, and
(**d**) proportional heteroscedasticity $\big[n = 99\,999,\ q(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}'\mathbb{P}$ where $\operatorname{vec}(\mathbb{P})' = \big(\exp(|Z|), 0, 0, \exp(-|Z|)\big)\big]$.

The four plots in Figure 1 illustrate all the core types of heteroscedastic behavior described in Section 5 with different numbers of observations. The elliptical regression $\tau$-$g_I$-quantiles, $\tau = 0.3$, were always computed from $n$ regression observations $(Y_1, Y_2, Z)$ generated by the regression model $(Y_1, Y_2)' = (Z, 0)' + q(\boldsymbol{\varepsilon})$ where $Z \sim U([-1,1])$, $\boldsymbol{\varepsilon} \sim U([-1,1]) \times U([-2,2])$ is independent of $Z$ (as everywhere below), and $q(\boldsymbol{\varepsilon})$ denotes a transformation of $\boldsymbol{\varepsilon}$ specific to each case. As for their parametrization by means of $\boldsymbol{s}_\tau$, $d_\tau$, $\alpha_\tau$, and $c_\tau$, always $\boldsymbol{s}_\tau = (\beta_1 Z, \beta_2)'$ and also $d_\tau^2 = \delta_1^2$, $\alpha_\tau = \alpha_1$, and $c_\tau = \gamma_1$ up to the exceptions listed below together with other specific features unique to individual pictures (a) to (d):

(**a**) no heteroscedasticity: $n = 99$, $q(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}$,

(**b**) rotational heteroscedasticity: $n = 999$, $\alpha_\tau = \pi\alpha_1 Z$, $q(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}'\mathbb{P}$ where $\operatorname{vec}(\mathbb{P})' = \big(\cos(\pi Z/2), \sin(\pi Z/2), -\sin(\pi Z/2), \cos(\pi Z/2)\big)$,

(**c**) multiplicative heteroscedasticity: $n = 9\,999$, $c_\tau = \gamma_1 + \gamma_2|Z| + \gamma_3 Z^2$, $q(\boldsymbol{\varepsilon}) = (0.1 + 0.9|Z|)\,\boldsymbol{\varepsilon}$, and

(**d**) proportional heteroscedasticity: $n = 99\,999$, $d_\tau^2 = \exp(\delta_1 Z)$, $q(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}'\mathbb{P}$ where $\operatorname{vec}(\mathbb{P})' = \big(\exp(|Z|), 0, 0, \exp(-|Z|)\big)$.

The objective function defining elliptical regression $\tau$-$g$-quantiles was optimized over all the scalar parameters occurring in the parametrization, as in all the following examples. In this case, it was over all $\boldsymbol{\theta} = (\beta_1, \beta_2, \delta_1, \alpha_1, \gamma_1, \gamma_2, \gamma_3)' \in \mathbb{R}^7$ in case (c) and over $\boldsymbol{\theta} = (\beta_1, \beta_2, \delta_1, \alpha_1, \gamma_1)' \in \mathbb{R}^5$ otherwise.

Figure 2 depicts elliptical regression $\tau$-$g_I$-quantiles with the trend, obtained for $\tau = 0.5$ from $n = 9\,999$ observations following the regression model $(Y_1, Y_2) = (Z, Z^2) + (1 + 3|\sin(\pi Z/2)|)\,\boldsymbol{\varepsilon}$ where $Z \sim U(-2, 2)$ and $\boldsymbol{\varepsilon} \sim N(0, 1/4) \times N(0, 1/4)$. They were parametrized with $\boldsymbol{s}_\tau = (\beta_1 + \beta_2 Z + \beta_3 Z^2, \beta_4 + \beta_5 Z + \beta_6 Z^2)'$, $d_\tau^2 = \delta_1^2$, $\alpha_\tau = \alpha_1$ and

(**a**) $c_\tau = \gamma_1$ or

(**b**) $c_\tau = \gamma_1 + \gamma_2|\sin(\pi Z/2)| + \gamma_3 \sin^2(\pi Z/2)$;

compare it to Figure 5 of [20] that is based on the same data generating model. This figure reminds you that one can easily enforce homoscedasticity or numerous equality constraints on model parameters when examining various submodels. In this particular case, the knowledge of the scale period is used in advance.

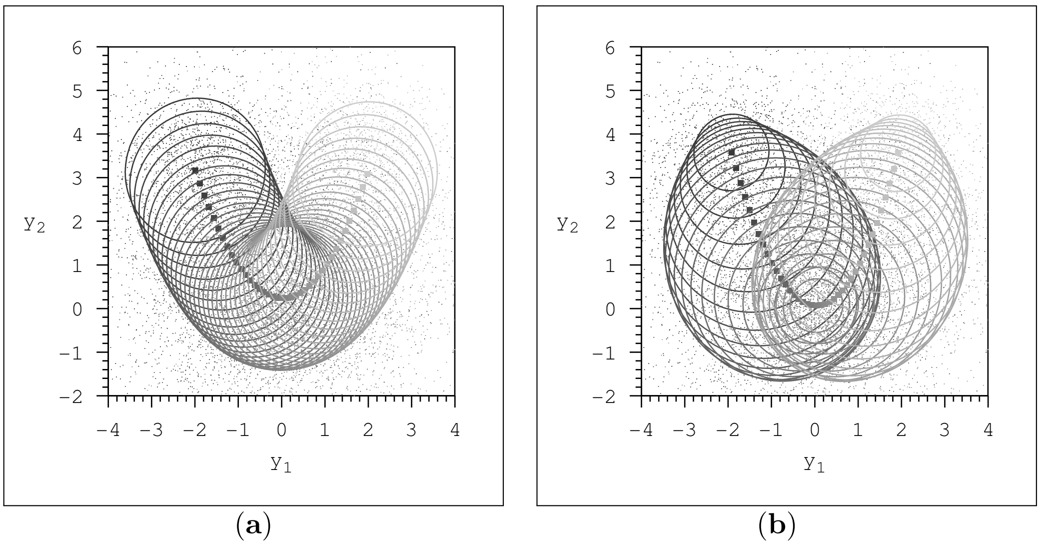


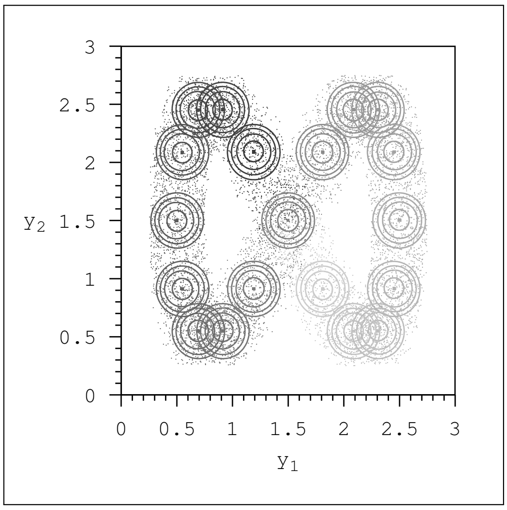

**Figure 2**: **Elliptical regression quantiles and a priori information.** The plots show elliptical regression $\tau$-$g_I$-quantile cuts and their centers, $\tau = 0.5$, obtained for reference points $z_0 = -1.9, -1.8, ..., 1.9$ from $n = 9\,999$ observations following the regression model $(Y_1, Y_2) = (Z, Z^2) + (1 + 3|\sin(\pi Z/2)|)\,\boldsymbol{\varepsilon}$ where $Z \sim U(-2, 2)$ is independent of $\boldsymbol{\varepsilon} \sim N(0, 1/4) \times N(0, 1/4)$. They assume a general quadratic trend in each component and
(**a**) homoscedasticity or
(**b**) the right form of heteroscedasticity.
Both the quantile curves and data points lighten with increasing values of the corresponding regressor.

Figure 3 is inspired by the well known Lissajous curves and highlights the fact that the parametric elliptical regression $\tau$-$g$-quantiles are especially convenient for fitting highly nonlinear models if one has an idea how to correctly describe the nonlinearity. They are computed for $\tau \in \{0.1, 0.3, 0.5, 0.7\}$ and $g_I$ from $n = 9\,999$ observations coming from a complicated nonlinear regression model $(Y_1, Y_2)' = \big(1.5 + \sin(Z), 1.5 + \sin(2Z)\big)' + q(\boldsymbol{\varepsilon}),\ Z \sim U([-\pi, \pi])$, $\boldsymbol{\varepsilon} \sim U([-0.25, 0.25]) \times U([-0.25, 0.25])$, where

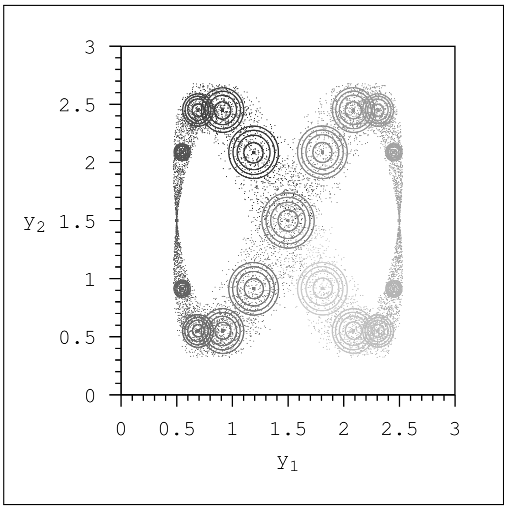
(**a**) $q(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}$ or

(**b**) $q(\boldsymbol{\varepsilon}) = \cos(Z)\,\boldsymbol{\varepsilon}$.

The quantile parameters were always looked for in the same form with generally $\tau$-dependent coefficients: $\boldsymbol{s}_\tau = \big(\beta_1 + \beta_2 \sin(\beta_3 Z), \beta_4 + \beta_5 \sin(\beta_6 Z)\big)'$, $d_\tau^2 = \delta_1^2$, $\alpha_\tau = \alpha_1$, and $c_\tau = \gamma_1 + \gamma_2^2 \cos^2(\gamma_3 Z)$.




(**a**)




(**b**)

**Figure 3**: **Elliptical regression quantiles and nonlinearity.** The plots display elliptical regression $\tau$-$g_I$-quantiles, $\tau \in \{0.1, 0.3, 0.5, 0.7\}$, for 19 equidistant reference points $z_0 = -9\pi/10, -8\pi/10, ..., 9\pi/10$, computed from $n = 9\,999$ observations coming from a complicated nonlinear regression model $(Y_1, Y_2)' = \big(1.5 + \sin(Z), 1.5 + \sin(2Z)\big)' + q(\boldsymbol{\varepsilon})$, $Z \sim U([-\pi, \pi])$ is independent of $\boldsymbol{\varepsilon} \sim U([-0.25, 0.25]) \times U([-0.25, 0.25])$, where
(**a**) $q(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}$ or
(**b**) $q(\boldsymbol{\varepsilon}) = \cos(Z)\,\boldsymbol{\varepsilon}$.
The quantile curves lighten with increasing $z_0$ and the data points get darker while the regressor values are decreasing.

The elliptical regression quantile methodology remains under investigation also in the next section where it is applied to real biometric data.

## 9. APPLICATION

For the sake of comparison, the parametric elliptical regression quantile methodology is tested on the same body girth measurements data of [19] as in [15], namely on $n = 260$ observations of calf maximum girth $Y_1$ (cm) and thigh (maximum) girth $Y_2$ (cm) of the physically active women whose age (years), weight (kg), height (cm) and body mass index (BMI $= 10\,000$ weight/height$^2$) are separately tried as the only regressor $Z$ in the attempts to explain $Y_1$ and $Y_2$. Although the observations do not constitute a random sample from any well-defined population, they are considered suitable for illustrating various statistical concepts.

In this particular case study, the parametric elliptical regression $\tau$-$g$-quantiles are computed for $g = g_I$. They are plotted only for $\tau \in \{0.1, 0.9\}$ and for $Z = z_0$ where $z_0$ is equal to the empirical $p$-th quantile of the regressor, $p \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$. The results are displayed in the same way as in Figure 7 of [15] to make the comparison as easy as possible. The only notable difference lies in the colors and quantile levels. That is to say that the pictures here are only black-and-white and, consequently, they illustrate the elliptical regression $\tau$-$g_I$-quantiles only for two representative values of $\tau$ to stay legible. Note also that the quantile levels used for indexing the elliptical regression quantiles by their overall probability coverage are not related to those used by the multiple-output directional quantile regression of [15] or [29] in any predictable way.

Figure 4 adopts the parametrization $\boldsymbol{s}_\tau = (\beta_1 + \beta_2 Z, \beta_3 + \beta_4 Z)'$, $d_\tau^2 = \delta_1^2$, $\alpha_\tau = \alpha_1$, and $c_\tau = \gamma_1 + \gamma_2 Z$ (with possibly different coefficients for each $\tau$) that allows for changes in location and scale and thus mimics the model used in [15] quite closely. Not surprisingly, it also produces similar output. Figures 4(a) and 4(c) clearly reveal certain location shift and scale increase of plotted $\tau$-quantile cuts caused by increasing weight and BMI, respectively. Figure 4(b) indicates that age influences only the location and volume of the outer quantile cuts but not of the inner ones. Figure 4(d) suggests that increasing height shifts both the inner and outer quantile cuts in mutually orthogonal directions but only affects the volume of the outer ones. Although all of these patterns can be more or less observed in Figure 7 of [15] as well, they are more clearly articulated through the simple elliptical shapes here. See also [14] and [17] for other quantile regression fits of the same data and their explanations.

Figure 5 plots the results regarding BMI for the generalized parametrization with $d_\tau^2 = (\delta_1 + \delta_2 Z)^2$, $\alpha_\tau = \alpha_1 + \alpha_2 Z$, and the other settings left unchanged, as in Figure 4(c). The modification permits more flexible changes of the regression quantile shape and is able to detect even the slight rotation of the outer quantile cuts with increasing BMI, observed in [15].

Although the analysis above is too simplistic to establish anything certain about female legs, it clearly demonstrates that the generalized parametric elliptical quantile regression is a powerful and flexible analytical method capable of pointing out even the smallest subtleties in the data behavior.

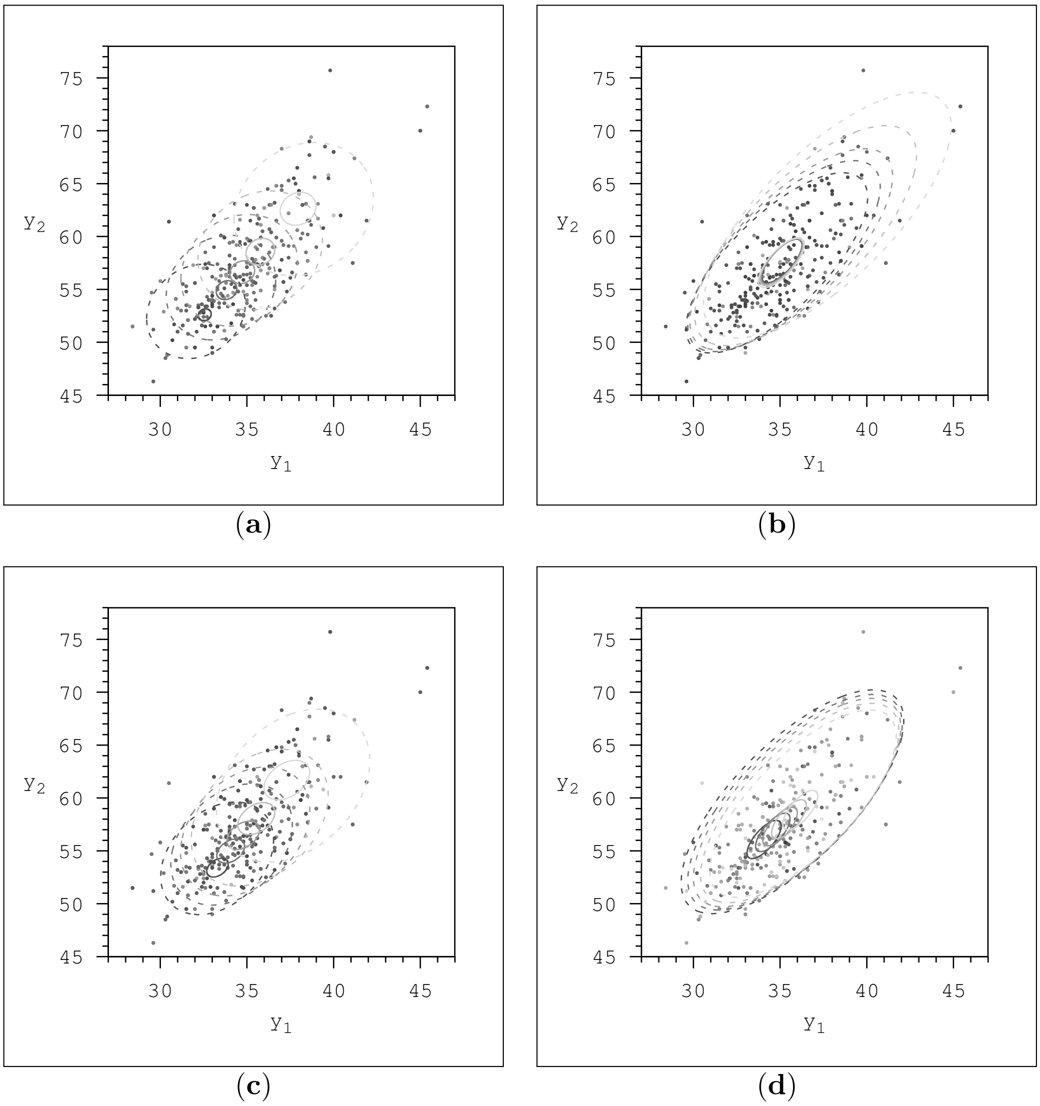



**Figure 4**: **Application to real data I.** The plots illustrate the dependence of female calf maximum girth ($Y_1$) and thigh (maximum) girth ($Y_2$) on
(**a**) weight,
(**b**) age,
(**c**) BMI, or
(**d**) height
by means of parametric elliptical $g_I$-quantile regression with a single regressor ($Z$), constant matrix parameter $\mathbb{A}$, linear inflation factor $c$, and linear trend $\boldsymbol{s}$. The elliptical regression $g_I$-quantiles are displayed for both $\tau = 0.1$ (solid line) and $\tau = 0.9$ (dashed line) and for regressor values $z_0$ equal to the empirical $p$-th quantile of $Z$, $p = 0.1, 0.3, 0.5, 0.7,$ and $0.9$. The quantile curves lighten with increasing $p$ and the data points get darker while the regressor values are decreasing.

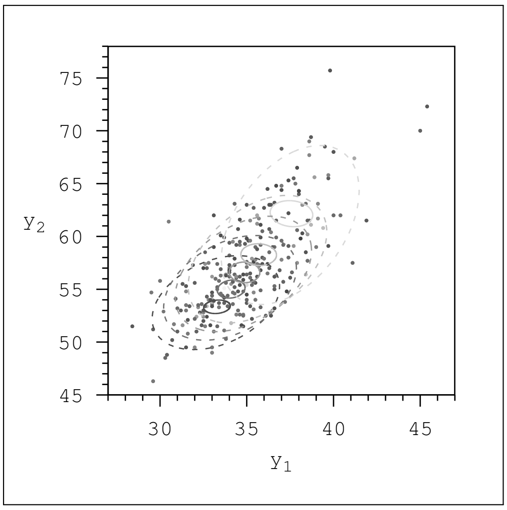



**Figure 5**: **Application to real data II.** The plot shows the dependence of female calf maximum girth ($Y_1$) and thigh (maximum) girth ($Y_2$) on BMI by means of parametric elliptical quantile regression assuming linear trend and a general form of heteroscedasticity. The elliptical regression $g_I$-quantiles are displayed for both $\tau = 0.1$ (solid line) and $\tau = 0.9$ (dashed line) and for regressor values $z_0$ equal to the empirical $p$-th quantile of the regressor, $p = 0.1, 0.3, 0.5, 0.7,$ and 0.9. The quantile curves lighten with increasing $p$ and the data points get darker while the regressor value is decreasing.

## 10. CONCLUDING REMARKS

All the presented theory and pictures demonstrate that the generalized parametric elliptical quantile regression may lead to natural and reasonable fits, even when the assumption of conditional symmetry cannot be relied on, as in Section 9. That is to say that the conditional central symmetry may simplify model validation and make the results from a well parametrized model particularly easy to interpret, but it is not strictly required for the method to work.

Sections 7, 8, and 9 also tacitly assume that the sample estimators of the quantile coefficients and cuts are consistent. It still has to be proved in full generality although it is already known in some special cases; see [17].

There is always a risk that the complicated non-convex optimization behind the generalized parametric elliptical quantile regression will terminate without finding the real global minimum. Nevertheless, this threat can be fought back by using global optimization strategies and model validation tools. And this problem should not theoretically appear at all for $g(t) = t$ and well-specified or specific models [17], and it is thus not likely to be severe in very similar situations.

The dependence of generalized parametric elliptical regression quantiles on function $g$ may rise another concern as it may seem to introduce too much arbitrariness into the model selection. However, simple fully affine equivariant parametrizations strongly ask for a power function $g$, and then its selection becomes as arbitrary as the choice of $p > 0$ in the standard $L_p$ regression. Only $L_2$ and $L_1$ regression methods are usually used because of their simplicity and easily interpretable results. And the same reasons lead to the choices $g(t) = t$ or $g(t) = \sqrt{t}$ in the generalized parametric elliptical quantile regression, though the latter seems reasonable only in certain special cases.

This article should be interpreted only as a single step on the long way to the successful elliptical quantile regression methodology. The next steps will include nonparametric generalizations, statistical inference, and a powerful and reliable software support.

It is difficult to predict if the proposed generalized parametric elliptical quantile regression withstands the test of time but, for the time being, it appears quite promising.

## ACKNOWLEDGMENTS

This research was supported by two Czech Science Foundation projects: GA14-07234S and GA17-07384S. Miroslav Šiman would like to thank Davy Paindaveine, Marc Hallin, Claude Adan, Nancy de Munck, and Romy Genin for their insight and encouragement, and also for all the good they did for him (and for all the good he could learn from them) during his stay at Université libre de Bruxelles.

## REFERENCES

[1] Ben, M.G.; Martínez, E. and Yohai, V.J. (2006). Robust estimation for the multivariate linear model based on a $\tau$-scale, *Journal of Multivariate Analysis*, **97**, 1600–1622.

[2] Boček, P. and Šiman, M. (2016). Directional quantile regression in Octave and MATLAB, *Kybernetika*, **52**, 28–51.

[3] Boček, P. and Šiman, M. (2017). Directional quantile regression in R, *Kybernetika*, **53**, 480–492.

[4] Boček, P. and Šiman, M. (2017). On weighted and locally polynomial directional quantile regression, *Computational Statistics*, **32**, 929–946.

[5] Carlier, G.; Chernozhukov, V. and Galichon, A. (2016). Vector quantile regression: an optimal transport approach, *Annals of Statistics*, **44**, 1165–1192.

[6] Chakraborty, B. (2003). On multivariate quantile regression, *Journal of Statistical Planning and Inference*, **110**, 109–132.

[7] Charlier, I.; Paindaveine, D. and Saracco, J. (2016). Multiple-output regression through optimal quantization, *ECARES Working Paper*, **2016-18**.

[8] Chaudhuri, P. (1996). On a geometric notion of quantiles for multivariate data, *Journal of the American Statistical Association*, **91**, 862–872.

[9] Cheng, Y. and De Gooijer, J.G. (2007). On the $u$-th geometric conditional quantile, *Journal of Statistical Planning and Inference*, **137**, 1914–1930.

[10] Chernozhukov, V.; Galichon, A.; Hallin, M. and Henry, M. (2017). Monge-Kantorovich depth, quantiles, ranks, and signs, *Annals of Statistics*, **45**, 223–256.

[11] De Gooijer, J.G.; Gannoun, A. and Zerom, D. (2006). A multivariate quantile predictor, *Communications in Statistics — Theory and Methods*, **35**, 133–147.

[12] Došlá, Š. (2009). Conditions for bimodality and multimodality of a mixture of two unimodal densities, *Kybernetika*, **45**, 279–292.

[13] Gómez, E.; Gómez-Villegas, M.A. and Marín, J.M. (2003). A survey on continuous elliptical vector distributions, *Revista Matemática Complutense*, **16**, 345–361.

[14] Hallin, M.; Lu, Z.; Paindaveine, D. and Šiman, M. (2015). Local bilinear multiple-output quantile/depth regression, *Bernoulli*, **21**, 1435–1466.

[15] Hallin, M.; Paindaveine, D. and Šiman, M. (2010). Multivariate quantiles and multiple-output regression quantiles: from $L_1$ optimization to halfspace depth, *Annals of Statistics*, **38**, 635–669.

[16] Hallin, M.; Paindaveine, D. and Šiman, M. (2010). Rejoinder, *Annals of Statistics*, **38**, 694–703.

[17] Hallin, M. and Šiman, M. (2016). Elliptical multiple-output quantile regression and convex optimization, *Statistics & Probability Letters*, **109**, 232–237.

[18] Hallin, M. and Šiman, M. (2017). *Multiple-output quantile regression*. In "Handbook of Quantile Regression" (R. Koenker, V. Chernozhukov, X. He and L. Peng, Eds.), Chapman and Hall/CRC.

[19] Heinz, G.; Peterson, L.J.; Johnson, R.W. and Kerjk, C.J. (2003). Exploring relationships in body dimensions, *Journal of Statistics Education*, **11**.

[20] Hlubinka, D. and Šiman, M. (2013). On elliptical quantiles in the quantile regression setup, *Journal of Multivariate Analysis*, **116**, 163–171.

[21] Hlubinka, D. and Šiman, M. (2015). On generalized elliptical quantiles in the nonlinear quantile regression setup, *TEST*, **24**, 249–264.

[22] Koenker, R. (2005). *Quantile Regression*, Cambridge University Press, New York.

[23] Koenker, R. and Bassett, G.J. (1978). Regression quantiles, *Econometrica*, **46**, 33–50.

[24] Koenker, R. and Park, B.J. (1996). An interior point algorithm for nonlinear quantile regression, *Journal of Econometrics*, **71**, 265–283.

[25] Koltchinskii, V. (1997). $M$-estimation, convexity and quantiles, *Annals of Statistics*, **25**, 435–477.

[26] Kong, L. and Mizera, I. (2012). Quantile tomography: using quantiles with multivariate data, *Statistica Sinica*, **22**, 1589–1610.

[27] McKeague, I.W.; López-Pintado, S.; Hallin, M. and Šiman, M. (2011). Analyzing growth trajectories, *Journal of Developmental Origins of Health and Disease*, **2**, 322–329.

[28] Paindaveine, D. (2008). A canonical definition of shape, *Statistics & Probability Letters*, **78**, 2240–2247.

[29] Paindaveine, D. and Šiman, M. (2011). On directional multiple-output quantile regression, *Journal of Multivariate Analysis*, **102**, 193–212.

[30] Paindaveine, D. and Šiman, M. (2012). Computing multiple-output regression quantile regions, *Computational Statistics & Data Analysis*, **56**, 840–853.

[31] Paindaveine, D. and Šiman, M. (2012). Computing multiple-output regression quantile regions from projection quantiles, *Computational Statistics*, **27**, 29–49.

[32] Rousseeuw, P.J.; Van Driessen, K.; Van Aelst, S. and Agulló, J. (2004). Robust multivariate regression, *Technometrics*, **46**, 293–305.

[33] SERFLING, R. (2002). Quantile functions for multivariate analysis: approaches and applications, *Statistica Neerlandica*, **56**, 214–232.

[34] ŠIMAN, M. (2011). On exact computation of some statistics based on projection pursuit in a general regression context, *Communications in Statistics — Simulation and Computation*, **40**, 948–956.

[35] ŠIMAN, M. (2014). Precision index in the multivariate context, *Communications in Statistics — Theory and Methods*, **43**, 377–387.

[36] VAN AELST, S. and WILLEMS, G. (2005). Multivariate regression $S$-estimators for robust estimation and inference, *Statistica Sinica*, **15**, 981–1001.

[37] 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.