



HAL
open science

ON EXACT COMPUTATION OF SOME STATISTICS BASED ON PROJECTION PURSUIT IN A GENERAL REGRESSION CONTEXT

Miroslav Šiman

► **To cite this version:**

Miroslav Šiman. ON EXACT COMPUTATION OF SOME STATISTICS BASED ON PROJECTION PURSUIT IN A GENERAL REGRESSION CONTEXT. *Communications in Statistics - Simulation and Computation*, 2011, 40 (06), pp.948-956. 10.1080/03610918.2011.560730 . hal-00681618

HAL Id: hal-00681618

<https://hal.science/hal-00681618v1>

Submitted on 22 Mar 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ON EXACT COMPUTATION OF SOME STATISTICS BASED ON PROJECTION PURSUIT IN A GENERAL REGRESSION CONTEXT

Journal:	<i>Communications in Statistics - Simulation and Computation</i>
Manuscript ID:	LSSP-2010-0268.R1
Manuscript Type:	Original Paper
Date Submitted by the Author:	26-Jan-2011
Complete List of Authors:	Šiman, Miroslav; Institute of Information Theory and Automation of the ASCR
Keywords:	exact computation, projection pursuit, quantile regression, multivariate quantile, projection depth, regression depth
Abstract:	It is shown in detail how recent advances in multiple-output and projectional quantile regression open the door to exact computation of many inferential statistics based on projection pursuit. This is also illustrated on a few examples including new regression generalizations of multivariate skewness, kurtosis and projection depth.
<p>Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.</p> <p>OnComputingSomeProjectionPursuitStatistics_AfterRevision.tex</p>	

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review Only

1 ON EXACT COMPUTATION OF SOME STATISTICS BASED ON PROJECTION PUR-
2 SUIT IN A GENERAL REGRESSION CONTEXT
3
4

5 Short title: ON COMPUTING SOME PROJECTION PURSUIT STATISTICS
6
7

8 Miroslav Šiman
9

10 Institute of Information Theory and Automation of the ASCR
11 Pod Vodárenskou věží 4, CZ-182 08, Prague 8, Czech Republic
12
13 siman@utia.cas.cz
14
15
16

17 Key Words: exact computation; projection pursuit; quantile regression; multivariate quan-
18 tile; projection depth; regression depth.
19
20
21

22 **ABSTRACT**

23 It is shown in detail how recent advances in multiple-output and projectional quantile regres-
24 sion open the door to exact computation of many inferential statistics based on projection
25 pursuit. This is also illustrated on a few examples including new regression generalizations
26 of multivariate skewness, kurtosis and projection depth.
27
28
29
30
31
32
33
34
35

36 **1 INTRODUCTION**

37 Recently, a new theory of directional (regression) quantiles has been developed, see Paindav-
38 eine and Šiman (2010b, 2010c) for the algorithms and Hallin et al. (2010a), Paindaveine and
39 Šiman (2010a), and Kong and Mizera (2008) for the theory. It gives rise to many promising
40 directional statistics but their directional dependence makes it difficult to exactly compute
41 their suprema or infima over infinitely many unit directions, which would be very useful
42 for overall statistical inference. This paper addresses the problem and solves it in some
43 important cases.
44
45
46
47
48
49
50
51

52 Besides, the projectional quantile regression of Paindaveine and Šiman (2010a, 2010c)
53
54

1 appears very useful for the exact computation of many statistics based on projection pursuit
2 methodology where the same problem arises as well; see e.g. Friedman (1987) and Huber
3 (1985). We illustrate this on a few important examples including multiple-output regression
4 generalizations of location projection depth, skewness and kurtosis. These concrete multi-
5 variate regression extensions are obtained quite naturally by means of regression quantiles,
6 introduced in Koenker and Bassett (1978) and masterly reviewed in Koenker (2005), but
7 they seem, quite surprisingly, not to have yet been explicitly considered in the literature.
8

9 The projection pursuit plays a crucial role in the definitions of various statistical depths
10 including halfspace depth, weighted halfspace depth and projection depth; see Rousseeuw
11 and Ruts (1999), Hlubinka et al. (2010) and Zuo (2003) with references therein, respectively.
12 The projection depth leads the way and beats many of its location depth competitors in nu-
13 merous respects, see Zuo and Serfling (2000a, 2000b). It has already been generalized even
14 to regression fits (i.e. *hyperplanes*) in Zuo and Cui (2004). Finally, it is here extended to
15 arbitrary *points* in a general multiple-output regression context, which seems more intuitive,
16 more useful and more promising for future development; see also Liu et al. (2004), Paindav-
17 eine and Šiman (2010a), and Rousseeuw and Hubert (1999) for other notions of depth in a
18 regression framework.
19

20 Exact computation of some special forms of location projection depth has already been
21 briefly outlined in Zuo (2004) for odd numbers of bivariate data points. Besides, there is
22 another paper dealing with exact computation of bivariate projection depth that is accom-
23 panied with software implementation in R, see Zuo (2010) and Zuo and Ye (2009). Although
24 that paper has not yet been published and made available, its title indicates that it deals
25 only with the bivariate and location case. In other words, the present paper almost surely
26 brings up some new ideas regarding this computational issue, at least in its treatment of
27 projection *regression* depth in spaces of *general* dimension. Needless to say that the results
28 presented here can lead to a real breakthrough in the dissemination of projection depth and
29 related methods.
30

31 Skewness and kurtosis of univariate (conditional) distributions can be described by means
32

of various statistics based on univariate (regression) quantiles, see e.g. Kim and White (2004) and White et al. (2008). Two of these univariate quantile-based shape characteristics are considered even here and extended by means of projection pursuit to the multiple-output context in the same way as their moment-based competitors in Malkovich and Afifi (1973). These extensions are then shown exactly computable even when the output dimension exceeds two or three, which may lead to many applications, including new tests for (conditional) multivariate symmetry or normality.

This paper proceeds as follows. Section 2 introduces one fundamental lemma that is then heavily applied to the multiple-output quantile regression in Section 3 and to the projectional quantile regression in Section 4. Its subsections discuss the exact computation of the multivariate shape statistics and projection depth in a general regression context. Section 5 concludes with some final remarks.

This brief communication blends so many various theoretical concepts together that their joint detail presentation in the Introduction would only confuse the reader. This is why their definitions are left only for the sections where they are employed. Sections 3 and 4 are independent, which explains why they sometimes use the same notation for similar but different entities.

2 THEORY

Hereafter, we adopt basic terminology of Boyd and Vandenberghe (2004).

Definition. *Function F on \mathbb{R}^m is directional (or, scale invariant) if $F(c\mathbf{u}) = F(\mathbf{u}) \forall c > 0 \forall \mathbf{u} \in \mathbb{R}^m$, and is sign-directional if $\text{sign}(F(c\mathbf{u})) = \text{sign}(F(\mathbf{u})) \forall c > 0 \forall \mathbf{u} \in \mathbb{R}^m$.*

Lemma 1. *Let us assume that the whole space \mathbb{R}^m , $m > 1$, can be partitioned into a finite number of non-degenerate closed convex polyhedral cones C_q with disjoint interiors, say $\mathbb{R}^m = \bigcup_{q=1}^N C_q$, and let us write $\Gamma := \{C_q : q = 1, \dots, N\}$ for this conic segmentation, determined by its finite vertex set $\mathcal{V} := \bigcup_{q=1}^N V_q$ where $V_q := \{\mathbf{u} \in \mathbb{R}^m : \|\mathbf{u}\| = 1, \mathbf{u} \text{ lies in an edge of } C_q\}$.*

1. *If F is directional and quasiconvex on each $C_q \in \Gamma$, then $\sup_{\|\mathbf{u}\|=1} F(\mathbf{u}) = \max_{\mathbf{u} \in \mathcal{V}} F(\mathbf{u})$.*

2. If F is directional and quasiconcave on each $C_q \in \Gamma$, then $\inf_{\|\mathbf{u}\|=1} F(\mathbf{u}) = \min_{\mathbf{u} \in \mathcal{V}} F(\mathbf{u})$.

3. If F is sign-directional and quasiconcave on each $C_q \in \Gamma$, then $F(\mathbf{u}) > 0$ for all $\mathbf{u} \in \mathcal{S}^{m-1} := \{\mathbf{u} \in \mathbb{R}^m : \|\mathbf{u}\| = 1\}$ if and only if $\min_{\mathbf{u} \in \mathcal{V}} F(\mathbf{u}) > 0$.

All these statements hold even when the assumptions are not met by F itself, but only by a function $\mathbf{u} \mapsto F(\mathbf{u})g(\|\mathbf{u}\|)$ where g is an arbitrary positive function on $(0, \infty)$ such as $g(x) = x$.

Proof. The surface D of the convex hull of \mathcal{V} consists of q closed convex polyhedral facets $D_q := \{\sum_i \lambda_i \mathbf{v}_i : \sum_i \lambda_i = 1, \lambda_i \geq 0, \mathbf{v}_i \in V_q\}$ determined by C_q , $D = \bigcup_{q=1}^N D_q$. If F is directional, then $\sup_{\|\mathbf{u}\|=1} F(\mathbf{u}) = \sup_{\mathbf{u} \in D} F(\mathbf{u}) = \max_{q=1, \dots, N} \sup_{\mathbf{u} \in D_q} F(\mathbf{u})$ and also $\inf_{\|\mathbf{u}\|=1} F(\mathbf{u}) = \inf_{\mathbf{u} \in D} F(\mathbf{u}) = \min_{q=1, \dots, N} \inf_{\mathbf{u} \in D_q} F(\mathbf{u})$. If F is quasiconvex on each C_q , then it is also quasiconvex on each D_q and $\sup_{\mathbf{u} \in D_q} F(\mathbf{u}) = \max_{\mathbf{u} \in V_q} F(\mathbf{u})$, see Boyd and Vandenberghe (2004). This proves the first statement.

If F is quasiconcave on each C_q , then $-F$ is quasiconvex there by definition, which implies $\inf_{\mathbf{u} \in D_q} F(\mathbf{u}) = \min_{\mathbf{u} \in V_q} F(\mathbf{u})$ and so $\inf_{\|\mathbf{u}\|=1} F(\mathbf{u}) = \min_{\mathbf{u} \in \mathcal{V}} F(\mathbf{u})$.

It also holds for any sign-directional function F that $F(\mathbf{u}) > 0$ on \mathcal{S}^{m-1} if and only if $F(\mathbf{u}) > 0$ on D , which is therefore equivalent to $\inf_{\mathbf{u} \in D} F(\mathbf{u}) = \min_{q=1, \dots, N} \min_{\mathbf{u} \in V_q} F(\mathbf{u}) > 0$ thanks to the quasiconcavity of F on each C_q .

Finally, $F(\mathbf{u})$ and $F(\mathbf{u})g(\|\mathbf{u}\|)$ differ on \mathcal{S}^{m-1} only by a positive multiplicative constant, which concludes the proof. \square

Lemma 1 will often be invoked below, but only with either $g(\|\mathbf{u}\|) = 1$ or $g(\|\mathbf{u}\|) = \|\mathbf{u}\|$ (convex in \mathbf{u}), which should always be kept in mind. The notation introduced in Lemma 1 will also be used hereinafter, together with the following definition.

Definition. Let us consider some indexed conic segmentations $\Gamma(i)$, $i \in I$, with their vertex sets $\mathcal{V}(i)$, $i \in I$. Then their common refinement $\Gamma(I)$ is defined as the conic segmentation determined by $\mathcal{V}(I) := \bigcup_{i \in I} \mathcal{V}(i)$.

All applications of Lemma 1 are likely to use the theory of quasiconvex functions, especially the following composition rules from Boyd and Vandenberghe (2004).

Note. It holds for arbitrary parameters \mathbb{A} , \mathbf{b} , \mathbf{c} and d of the right dimensions that

1. if $g : \mathbb{R}^m \rightarrow \mathbb{R}$ is quasiconvex and $h : \mathbb{R} \rightarrow \mathbb{R}$ is nondecreasing, then $f = h \circ g$ is quasiconvex
2. if g is quasiconvex, then $f(\mathbf{u}) := g(\mathbb{A}\mathbf{u} + \mathbf{b})$ is quasiconvex
3. if g is quasiconvex, then $f(\mathbf{u}) := g((\mathbb{A}\mathbf{u} + \mathbf{b})/(\mathbf{c}'\mathbf{u} + d))$ is quasiconvex on the set $\{\mathbf{u} \in \mathbb{R}^m : \mathbf{c}'\mathbf{u} + d > 0\}$
4. if g is a convex function, h is a concave function, and both $g(\mathbf{u}) > 0$ and $h(\mathbf{u}) > 0$ on a convex set C , then $f(\mathbf{u}) = g(\mathbf{u})/h(\mathbf{u})$ is quasiconvex on C

where all the compositions are well-defined by assumption.

For the sake of simplicity, only suprema of directional statistics will be discussed hereinafter (without any significant loss of generality). This is why these composition rules are formulated only for quasiconvex functions here. They can be used to check or extend all the claims made below.

Let us consider m -dimensional responses \mathbf{Y}_i associated with $(p+1)$ -dimensional regressors $\mathbf{X}_i = (1, \mathbf{Z}_i)'$, $i = 1, \dots, n > p + 2$, and assume that the random sample $(\mathbf{Y}_i, \mathbf{Z}_i) \in \mathbb{R}^{m+p}$, $i = 1, \dots, n$, comes from a continuous distribution. **Section 3 and 4 discuss the consequences of Lemma 1 in the empirical regression context with a suitable conic segmentation.**

3 MULTIPLE-OUTPUT QUANTILE REGRESSION

Hallin et al. (2010a) introduced a new concept of directional multiple-output regression quantiles. In the general empirical case, they can be defined for any $\mathbf{u} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ by means of the optimization problem

$$(\mathbf{a}'_{\tau\mathbf{u}}, \mathbf{b}'_{\tau\mathbf{u}})' = \underset{(\mathbf{a}', \mathbf{b}')'}{\operatorname{argmin}} \sum_{i=1}^n w_i \rho_{\tau}(\mathbf{b}'\mathbf{Y}_i - \mathbf{a}'\mathbf{X}_i) \quad \text{subject to} \quad \mathbf{u}'\mathbf{b} = 1 \quad (1)$$

where $w_i > 0, i = 1, \dots, n$, are some positive weights, $\tau \in (0, 1)$ stands for any fixed quantile level, $\rho_\tau(x) = x(\tau - I(x < 0))$ is the well-known quantile check function and

$$\Psi_{\tau\mathbf{u}} = \sum_{i=1}^n w_i \rho_\tau(\mathbf{b}'_{\tau\mathbf{u}} \mathbf{Y}_i - \mathbf{a}'_{\tau\mathbf{u}} \mathbf{X}_i) \quad (2)$$

denotes the optimal value of the objective function. Typically, $w_i = 1, i = 1, \dots, n$, as in Hallin et al. (2010a).

It follows from Paindaveine and Šiman (2010b) that there exists a finite conic segmentation $\Gamma(\tau) = \{C_q(\tau) : q = 1, \dots, N_\tau\}$ of \mathbb{R}^m almost surely for all but a finite number of τ 's such that it meets the assumptions of Lemma 1 and $(\mathbf{a}'_{\tau\mathbf{u}}, \mathbf{b}'_{\tau\mathbf{u}})' = (\mathbf{a}'_q, \mathbf{b}'_q)' / \mathbf{t}'_q \mathbf{u}$, $\Psi_{\tau\mathbf{u}} = \psi_q^\tau / \mathbf{t}'_q \mathbf{u}$ and $\mathbf{t}'_q \mathbf{u} > 0$ for any \mathbf{u} in any $C_q(\tau)$ where $\mathbf{a}_q^\tau, \mathbf{b}_q^\tau, \mathbf{t}_q^\tau$ and ψ_q^τ are some constant vectors or scalars possibly different on each $C_q(\tau)$. Paindaveine and Šiman (2010b) also provide an algorithm (and its MATLAB implementation by means of parametric programming) that can find such conic segmentation $\Gamma(\tau)$, its vertex set $\mathcal{V}(\tau)$ and corresponding characteristics $\mathbf{a}_q^\tau, \mathbf{b}_q^\tau, \mathbf{t}_q^\tau$ and $\psi_q^\tau, q = 1, \dots, N_\tau$.

Clearly, $(\mathbf{a}'_{\tau\mathbf{u}}, \mathbf{b}'_{\tau\mathbf{u}})' = -(\mathbf{a}'_{(1-\tau)(-\mathbf{u})}, \mathbf{b}'_{(1-\tau)(-\mathbf{u})})'$ and thus $\mathcal{V}(\tau) = -\mathcal{V}(1 - \tau)$. Note as well that $\Gamma(T)$ is easy to construct for any finite index set T of τ 's, at least for $m = 2$, and that $\mathbf{a}_{\tau\mathbf{u}}, \mathbf{b}_{\tau\mathbf{u}}$ and $\Psi_{\tau\mathbf{u}}$ can be computed for any $\mathbf{u} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ by means of available $\mathbf{a}_q^\tau, \mathbf{b}_q^\tau, \mathbf{t}_q^\tau$ and ψ_q^τ , which may be more effective than their straightforward computation from scratch.

3.1 Simple Inference in Multiple-Output Quantile Regression

Hallin et al. (2010a) already showed that both $\mathbf{a}_{\tau\mathbf{u}}, \mathbf{b}_{\tau\mathbf{u}}$ and $\Psi_{\tau\mathbf{u}}$ can be very useful for statistical inference and that they can be combined into directional statistics that are constant on each $C_q(\tau)$ and therefore easy to maximize or minimize over \mathcal{S}^{m-1} .

Lemma 1 can be applied in this context to extend these results considerably. For example, it implies $\sup_{\|\mathbf{u}\|=1} G(\mathbf{u}) = \max_{\mathbf{u} \in \mathcal{V}(\tau)} G(\mathbf{u})$ where $G(\mathbf{u})$ can stand e.g. for $\|\mathbf{a}_{\tau\mathbf{u}}\|, \|\mathbf{b}_{\tau\mathbf{u}}\|$ and $\Psi_{\tau\mathbf{u}}$. Similarly, it guarantees for any suitable $\tau_1, \tau_2 \in (0, 1)$ that $\Psi_{\tau_1\mathbf{u}} / \Psi_{\tau_2\mathbf{u}}$ has its supremum over \mathcal{S}^{m-1} respectively equal to its maximum over the vertex set $\mathcal{V}(T)$ corresponding to $\Gamma(T), T = \{\tau_1, \tau_2\}$.

4 PROJECTIONAL QUANTILE REGRESSION

Alternatively, one can consider projectional regression quantiles from Kong and Mizera (2008) and Paindaveine and Šiman (2010a). In the general empirical case, they can be defined for any $\mathbf{u} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ by means of the optimization problem

$$\mathbf{a}_{\tau\mathbf{u}} := \operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^{p+1}} \sum_{i=1}^n w_i \rho_{\tau}(\mathbf{u}'\mathbf{Y}_i - \mathbf{a}'\mathbf{X}_i) \quad (3)$$

where $w_i > 0, i = 1, \dots, n$, are some positive weights, $\tau \in (0, 1)$ denotes the quantile level, $\rho_{\tau}(x) = x(\tau - \mathbf{I}(x < 0))$ is the well-known quantile check function and

$$\Psi_{\tau\mathbf{u}} := \sum_{i=1}^n w_i \rho_{\tau}(\mathbf{u}'\mathbf{Y}_i - \mathbf{a}'_{\tau\mathbf{u}}\mathbf{X}_i) \quad (4)$$

stands for the optimal value of the objective function. Often $w_i = 1, i = 1, \dots, n$, and then $\mathbf{a}_{\tau\mathbf{u}}$ is nothing but the standard sample regression τ -quantile of projections $\mathbf{u}'\mathbf{Y}_i$'s.

It was shown in Paindaveine and Šiman (2010a, 2010c) that there exists a finite conic segmentation $\Gamma(\tau) = \{C_q : q = 1, \dots, N_{\tau}\}$ of \mathbb{R}^m almost surely for all but a finite number of τ 's such that each C_q is a non-degenerate closed convex polyhedral cone where $\mathbf{a}_{\tau\mathbf{u}}$ and $\Psi_{\tau\mathbf{u}}$ are linear in \mathbf{u} , i.e. $\mathbf{a}_{\tau\mathbf{u}} = \mathbb{A}_q^{\tau}\mathbf{u}, \mathbb{A}_q^{\tau} \in \mathbb{R}_{(p+1) \times m}$, and $\Psi_{\tau\mathbf{u}} = \boldsymbol{\psi}_q^{\tau'}\mathbf{u}, \boldsymbol{\psi}_q^{\tau} \in \mathbb{R}^m$, for any $\mathbf{u} \in C_q(\tau)$. Paindaveine and Šiman (2010c) also provide an algorithm (and its MATLAB implementation by means of parametric programming) that can find such a conic segmentation $\Gamma(\tau)$, its vertex set $\mathcal{V}(\tau)$ and corresponding matrices \mathbb{A}_q^{τ} and vectors $\boldsymbol{\psi}_q^{\tau}, q = 1, \dots, N_{\tau}$.

Clearly, $\mathbf{a}_{\tau\mathbf{u}} = -\mathbf{a}_{(1-\tau)(-\mathbf{u})}$ and thus $\mathcal{V}(\tau) = -\mathcal{V}(1-\tau)$. Note as well that $\Gamma(T)$ is easy to construct for any index set T of τ 's, at least for $m = 2$, and that $\mathbf{a}_{\tau\mathbf{u}}$ and $\Psi_{\tau\mathbf{u}}$ can be computed for any $\mathbf{u} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ by means of available \mathbb{A}_q^{τ} and $\boldsymbol{\psi}_q^{\tau}$, which may be more effective than their straightforward computation from scratch.

4.1 Simple Inference in Projectional Quantile Regression

Paindaveine and Šiman (2010a) already showed that both $\mathbf{a}_{\tau\mathbf{u}}$ and $\Psi_{\tau\mathbf{u}}$ can be very useful for statistical inference. Now Lemma 1 makes such inference possible and simple in many

cases. For example, its direct application guarantees for any suitable $\tau_1, \tau_2 \in (0, 1)$ that $\sup_{\|\mathbf{u}\|=1} G(\mathbf{u}) = \max_{\mathbf{u} \in \mathcal{V}(\tau_1) \cup \mathcal{V}(\tau_2)} G(\mathbf{u})$ where $G(\mathbf{u})$ can stand for $\Psi_{\tau_1 \mathbf{u}} / \Psi_{\tau_2 \mathbf{u}}$, $\|\mathbf{a}_{\tau_1 \mathbf{u}}\| / \Psi_{\tau_2 \mathbf{u}}$, or $1 / \Psi_{\tau_1 \mathbf{u}}$, among others.

4.2 Ratios and Inverses of Regression L-Statistics

Lemma 1 can also be applied to ratios $R_{\mathbf{u}}(\mathbf{x})$ of directional regression L-statistics:

$$R_{\mathbf{u}}(\mathbf{x}) = \frac{\sum_{\tau \in T_N} c_{\tau} \mathbf{a}'_{\tau \mathbf{u}} \mathbf{x}}{\sum_{\tau \in T_D} d_{\tau} \mathbf{a}'_{\tau \mathbf{u}} \mathbf{x}} := \frac{N_{\mathbf{u}}(\mathbf{x})}{D_{\mathbf{u}}(\mathbf{x})}.$$

The denominator $D_{\mathbf{u}}(\mathbf{x})$ is linear in \mathbf{u} on each cone from $\Gamma(T_D)$ and therefore it is positive for all $\mathbf{u} \in \mathcal{S}^{m-1}$ if and only if its minimum over $\mathcal{V}(T_D)$ is greater than zero, which can be easily checked. If this is the case, then $\sup_{\|\mathbf{u}\|=1} R_{\mathbf{u}}(\mathbf{x})$ can be computed as $\max_{\mathbf{u} \in \mathcal{V}(T_D \cup T_N)} R_{\mathbf{u}}(\mathbf{x})$ because $R_{\mathbf{u}}(\mathbf{x})$ is then even quasilinear in \mathbf{u} on each cone from $\Gamma(T_D \cup T_N)$.

Lemma 1 also implies $\sup_{\mathbf{u} \in \mathcal{S}} 1/D_{\mathbf{u}}(\mathbf{x}) = \max_{\mathbf{u} \in \mathcal{V}(T_D)} 1/D_{\mathbf{u}}(\mathbf{x})$ if $D_{\mathbf{u}}(\mathbf{x}) > 0 \forall \mathbf{u} \in \mathcal{S}^{m-1}$.

For example, given $\mathbf{x} \in 1 \times \mathbb{R}^p$, we can consider two successful quantile-based shape measures from Kim and White (2006) adjusted to the general directional regression framework considered here, namely the skewness coefficient $\xi_{\mathbf{u}}(\mathbf{x})$ and the kurtosis measure $\kappa_{\mathbf{u}}(\mathbf{x})$:

$$\xi_{\mathbf{u}}(\mathbf{x}) = \frac{(\mathbf{a}'_{(3/4)\mathbf{u}} - \mathbf{a}'_{(2/4)\mathbf{u}}) \mathbf{x} - (\mathbf{a}'_{(2/4)\mathbf{u}} - \mathbf{a}'_{(1/4)\mathbf{u}}) \mathbf{x}}{(\mathbf{a}'_{(3/4)\mathbf{u}} - \mathbf{a}'_{(1/4)\mathbf{u}}) \mathbf{x}},$$

$$\kappa_{\mathbf{u}}(\mathbf{x}) = \frac{(\mathbf{a}'_{(7/8)\mathbf{u}} - \mathbf{a}'_{(5/8)\mathbf{u}}) \mathbf{x} + (\mathbf{a}'_{(3/8)\mathbf{u}} - \mathbf{a}'_{(1/8)\mathbf{u}}) \mathbf{x}}{(\mathbf{a}'_{(6/8)\mathbf{u}} - \mathbf{a}'_{(2/8)\mathbf{u}}) \mathbf{x}}.$$

Both these statistics can be rewritten in the form of $R_{\mathbf{u}}(\mathbf{x})$. Indeed, the former would lead to $T_D(\xi) = \{1/4, 3/4\}$ and $T_N(\xi) = \{1/4, 2/4, 3/4\}$ while the latter would correspond to $T_D(\kappa) = \{2/8, 6/8\}$ and $T_N(\kappa) = \{1/8, 3/8, 5/8, 7/8\}$. If their denominators are always positive, which holds almost trivially in the location case with $p = 0$, then we can meaningfully define overall skewness $\xi(\mathbf{x}) := \sup_{\mathbf{u} \in \mathcal{S}} \xi_{\mathbf{u}}(\mathbf{x})$ and overall kurtosis $\kappa(\mathbf{x}) := \sup_{\mathbf{u} \in \mathcal{S}} \kappa_{\mathbf{u}}(\mathbf{x})$ by means of projection pursuit and compute them as $\max_{\mathbf{u} \in \mathcal{V}(T_D(\xi) \cup T_N(\xi))} \xi_{\mathbf{u}}(\mathbf{x})$ and $\max_{\mathbf{u} \in \mathcal{V}(T_D(\kappa) \cup T_N(\kappa))} \kappa_{\mathbf{u}}(\mathbf{x})$, respectively.

4.3 Projection (Regression) Depth

Let us generalize location projection depth, thoroughly investigated in Zuo (2003), to the points in the general regression setup. In the empirical case discussed here, the projection regression depth $PD_n(\mathbf{y}, \mathbf{x})$ of $\mathbf{y} \in \mathbb{R}^m$ given $\mathbf{x} = (1, \mathbf{z}')' \in 1 \times \mathbb{R}^p$ may be defined as

$$PD_n(\mathbf{y}, \mathbf{x}) = \frac{1}{1 + \sup_{\|\mathbf{u}\|=1} g_n(\mathbf{y}, \mathbf{x}, \mathbf{u})}$$

with

$$g_n(\mathbf{y}, \mathbf{x}, \mathbf{u}) = \frac{|\mathbf{u}'\mathbf{y} - \mu_{\mathbf{u}}(\mathbf{x})|}{\sigma_{\mathbf{u}}(\mathbf{x})}$$

where $\mu_{\mathbf{u}}(\mathbf{x})$ and $\sigma_{\mathbf{u}}(\mathbf{x}) > 0$ are some sample suitable regression measures of location and scatter of $\mathbf{u}'\mathbf{Y}$ given \mathbf{X} , respectively.

In this example, let us consider only $\mu_{\mathbf{u}}(\mathbf{x})$ and $\sigma_{\mathbf{u}}(\mathbf{x})$ in the form of regression L -statistics,

$$\mu_{\mathbf{u}}(\mathbf{x}) := \sum_{\tau \in T_{\mu}} c_{\tau} \mathbf{a}'_{\tau} \mathbf{x} \quad \text{and} \quad \sigma_{\mathbf{u}}(\mathbf{x}) := \sum_{\tau \in T_{\sigma}} d_{\tau} \mathbf{a}'_{\tau} \mathbf{x},$$

where $c_{\tau} \in \mathbb{R}$, $\tau \in T_{\mu}$, and $d_{\tau} \in \mathbb{R}$, $\tau \in T_{\sigma}$. This restriction does not seem too severe after considering the close asymptotic relationship between R-, M-, and L-statistics thoroughly described in Jurečková and Sen (1996).

Generally speaking, the projection (regression) depth is difficult to compute fast and exactly for $m > 1$ because the supremum in its definition then involves infinitely many unit directions. Besides, its regression definition seems reasonable only if $\sigma_{\mathbf{u}}(\mathbf{x}) > 0$ for any $\mathbf{u} \in \mathcal{S}^{m-1}$, which need not be true for $p > 0$ due to the quantile crossing phenomenon.

Nevertheless, Lemma 1 implies that $\sigma_{\mathbf{u}}(\mathbf{x}) > 0$ for any $\mathbf{u} \in \mathcal{S}^{m-1}$ if and only if $\min_{\mathbf{u} \in \mathcal{V}(T_{\sigma})} \sigma_{\mathbf{u}}(\mathbf{x}) > 0$, which can be easily checked. And if this condition holds, then $g_n(\mathbf{y}, \mathbf{x}, \mathbf{u})$ is directional and quasiconvex on each $C_q \in \Gamma(T_{\mu} \cup T_{\sigma})$, which further results in

$$PD_n(\mathbf{y}, \mathbf{x}) = \frac{1}{1 + \max_{\mathbf{u} \in \mathcal{V}(T_{\mu} \cup T_{\sigma})} g_n(\mathbf{y}, \mathbf{x}, \mathbf{u})}.$$

Note that this formula can be used for fast and exact computation of projection depth even in the general regression context and for $m > 2$. The choice of T_{μ} and T_{σ} allows for a

1 compromise between robustness and speed as the computational time of finding $\Gamma(\tau)$ and
2 related quantities grows with decreasing $|\tau - 0.5|$, see Paindaveine and Šiman (2010c).
3
4
5
6

7 5 FINAL REMARKS

8
9
10 The minimizers $(\mathbf{a}'_{\tau\mathbf{u}}, \mathbf{b}'_{\tau\mathbf{u}})'$ in Section 3 and $\mathbf{a}_{\tau\mathbf{u}}$ in Section 4 need not be defined uniquely, not
11 even in the purely location setup for integer values of $n\tau$, but there is only a finite number
12 of such exceptional quantile levels with probability one. Besides, this potential ambiguity
13 is not too worrying because all such minimizers always form a convex polyhedral set that
14 shrinks towards their unique population limits under very mild conditions, see Hallin et
15 al. (2010a) and Chapter 4 in Koenker (2005) for a detailed discussion of these asymptotic
16 properties. Anyway, such an ambiguity can often be fixed easily by a tiny perturbation of τ
17 or the data points, which can have hardly any impact on practical applications. This is why
18 these minimizers are considered uniquely defined throughout this paper.
19
20
21
22
23
24
25
26

27 Clearly, the theory presented here is especially useful when the number of cones in the
28 conic segmentations is quite low. In Sections 3 and 4, this often happens for small n or
29 extreme τ . Then even the search for approximate extremes on a sampled dense grid of
30 directions would be much more complicated, in particular for $m > 2$.
31
32
33
34

35 It should also be pointed out that the weights w_i 's used in Sections 3 and 4 can be really
36 useful, both for handling multiple observations and for computing local constant counterparts
37 to the regression quantiles of Sections 3 and 4, see Hallin et al. (2010b). They would also lead
38 to interesting local versions of the shape measures and projection regression depth considered
39 in Section 4.
40
41
42
43
44
45

46 ACKNOWLEDGEMENTS

47
48
49 This research work of Miroslav Šiman was supported by Project 1M06047 of the Ministry of
50 Education, Youth and Sports of the Czech Republic. The author would also like to thank
51
52
53
54

1 Davy Paindaveine, Marc Hallin, Claude Adan, Nancy de Munck and Romy Genin for all the
2 good they did for him (and for all the good he could learn from them) during his recent stay
3 at Université Libre de Bruxelles.
4
5

6 BIBLIOGRAPHY

7
8
9
10 Boyd, S., and Vandenberghe, L. (2004). *Convex optimization*. Cambridge: Cambridge
11 University Press. http://www.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf.

12
13
14 Friedman, J.H. (1987). Exploratory projection pursuit. *Journal of the American Statistical*
15 *Association*, **82**, 249–266.
16

17
18
19 Hallin, M., Paindaveine, D., and Šiman, M. (2010a). Multivariate quantiles and multiple-
20 output regression quantiles: from L_1 optimization to halfspace depth. *Annals of Statistics*,
21 **38**, 635–669.
22

23
24
25 Hallin, M., Lu, Z., Paindaveine, D., and Šiman, M. (2010b). Local bilinear multiple-output
26 quantile regression and regression depth. Manuscript in preparation.
27

28
29
30 Hlubinka, D., Kotík, L., and Vencálek, O. (2010). Weighted halfspace depth. *Kybernetika*,
31 **46**, 125–148.
32

33
34
35 Huber, P.J. (1985). Projection pursuit. *Annals of Statistics*, **13**, 435–475.
36

37
38
39 Jurečková, J., and Sen, P.K. (1996). *Robust statistical procedures: asymptotics and interre-*
40 *lations*. New York: Wiley.
41

42
43
44 Kim, T.-H., and White, H. (2004). On more robust estimation of skewness and kurtosis.
45 *Finance Research Letters*, **1**, 56–73.
46

47
48
49 Koenker, R., and Basset, G.J. (1978). Regression quantiles. *Econometrica*, **46**, 33–50.
50

51
52
53 Koenker, R. (2005). *Quantile Regression*. New York: Cambridge University Press.
54

55
56
57 Kong, L., and Mizera, I. (2008). Quantile tomography: using quantiles with multivariate
58
59
60

1 data. arXiv:0805.0056v1.

2
3
4 Liu, R.Y., Singh, K., and Teng, J.H. (2004). Linear fitting by simplicial intercept depth
5 (SID): reflection invariance and robustness. *Statistica Sinica*, **14**, 431–448.

6
7
8 Malkovich, J.F., and Afifi, A.A. (1973). On tests of multivariate normality. *Journal of the*
9 *American Statistical Association*, **68**, 176–179.

10
11
12 Paindaveine, D., and Šiman, M. (2010a). On directional multiple-output quantile regression.
13 *Journal of Multivariate Analysis*, accepted.

14
15
16 Paindaveine, D., and Šiman, M. (2010b). Computing multiple-output regression quantile
17 regions. *Computational Statistics & Data Analysis*, tentatively accepted.

18
19
20 Paindaveine, D., and Šiman, M. (2010c). Computing multiple-output regression quantile
21 regions from projection quantiles. *Computational Statistics*, tentatively accepted.

22
23
24 Rousseeuw, P.J., and Hubert, M. (1999). Regression depth. *Journal of the American Sta-*
25 *tistical Association*, **94**, 388–402.

26
27
28 Rousseeuw, P.J., and Ruts, I. (1999). The depth function of a population distribution.
29 *Metrika*, **49**, 213–244.

30
31
32 White, H., Kim, T.-H., and Manganelli, S. (2008). Modeling autoregressive conditional
33 skewness and kurtosis with multi-quantile CAViaR. ECB working paper.

34
35
36 Zuo, Y. (2003). Projection-based depth functions and associated medians. *Annals of Statis-*
37 *tics*, **31**, 1460–1490.

38
39
40 Zuo, Y. (2004). Projection-based affine equivariant multivariate location estimators with the
41 best possible finite sample breakdown point. *Statistica Sinica*, **14**, 1199–1208.

42
43
44 Zuo, Y. (2010). Exact computation of the bivariate projection depth and Stahel-Donoho
45 estimator. *Computational Statistics & Data Analysis*, tentatively accepted.

1 Zuo, Y., and Cui, H. (2004). Statistical depth functions and some applications. *Advances*
2 *in Mathematics*, **33**, 1–25.

3
4
5 Zuo, Y., and Serfling, R. (2000a). On the performance of some robust nonparametric loca-
6 tion measures relative to a general notion of multivariate symmetry. *Journal of Statistical*
7 *Planning and Inference*, **84**, 55–79.

8
9
10 Zuo, Y., and Serfling, R. (2000b). General notions of statistical depth functions. *Annals of*
11 *Statistics*, **28**, 461–482.

12
13
14 Zuo, Y., and Ye, X. (2009). ExPD2D: Exact computation of bivariate projection depth based
15 on Fortran code. R package version 1.0.1, <http://cran.r-project.org/package=ExPD2D>.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60