

Target Asymptotic Problems Arising with Multivariate Sample Quantile Functions

Robert Serfling¹

University of Texas at Dallas

New Directions in Asymptotic Statistics Symposium
May 15, 2009 – updated June 14, 2009

¹www.utdallas.edu/~serfling

Depth, Outlyingness, Quantile, and Rank Functions in \mathbb{R}^d

Definitions and Examples

Equivalence: the D-O-Q-R Paradigm

Equivariance/Invariance Properties

Functionals for Standardization, and Their Roles

A Simple Message

Practical Applications Using D-O-Q-R Approaches

Diversity and Complexity of the D-O-Q-R Landscape

Computational Burden & Computational Geometry

Desirable Results for Sample Versions

Asymptotic Results for D-O-Q-R Methods: A Brief Sketch

Invitation and Challenge

A Few Concluding Remarks

Multivariate *Depth, Outlyingness, Quantiles, Ranks*

- ▶ “*Order statistics*”, “*outlier identification*”, “*quantiles*”, “*signs*”, and “*ranks*” have their own special roles and their own “practitioners” and “afficionados”.
- ▶ Along with “*symmetry*”, these together comprise the *fundamental elements of nonparametric description*.
- ▶ *Intuitively*, they are *interrelated*.
- ▶ In fact, D, O, Q, and R are equivalent methodologies.
- ▶ We “prove” this claim by making *the right definitions!*

Outlyingness Functions: *Finding “Outliers”*

- ▶ “Outliers” have been under discussion for centuries:
Francis Bacon, 1620 ... Daniel Bernoulli, 1777 ...
Benjamin Pierce, 1852 ... Barnett and Lewis, 1995 ...
- ▶ In *higher dimension*, detection by visualization fails.
- ▶ Thus we need algorithmic approaches:

Given a cdf F on \mathbb{R}^d , an outlyingness function $O(\mathbf{x}, F)$ is an associated *center-outward ordering* of points \mathbf{x} in \mathbb{R}^d with *higher* values representing greater “outlyingness”.

Finding Outliers via SPSS (For Example)

And SPSS offers “algorithms” for you!!!

SPSS’s advertisement in *Amstat News*, 2007:

Quickly find multivariate outliers

Prevent outliers from skewing analyses when you use the Anomaly Detection Procedure. This procedure searches for unusual cases based on deviations from similar cases and gives reasons for such deviations. You can flag outliers by creating a new variable. Once you have identified unusual cases, you can further examine them and determine if they should be included in your analyses.

Example: *Projection Pursuit Approach*

- ▶ Given a *univariate* outlyingness function $O_1(\cdot, \cdot)$, define a multivariate extension by

$$O_d(\mathbf{x}, F) = \sup_{\|\mathbf{u}\|=1} O_1(\mathbf{u}'\mathbf{x}, F_{\mathbf{u}'\mathbf{x}}), \mathbf{x} \in \mathbb{R}^d$$

- ▶ For example, with *univariate scaled deviation*

$$O_1(x, F) = \left| \frac{x - \mu(F)}{\sigma(F)} \right|,$$

for $\mu(F)$ and $\sigma(F)$ location and spread measures, e.g., $\text{Med}(F)$ and $\text{MAD}(F)$, this $O_d(\mathbf{x}, F)$ is *affine invariant*.

Example: *Spatial Sign Function Approach*

- ▶ Using the d -dimensional 'spatial' sign function (or *unit vector function*),

$$\mathbf{S}(\mathbf{x}) = \frac{\mathbf{x}}{\|\mathbf{x}\|}, \quad \mathbf{x} \in \mathbb{R}^d,$$

we obtain the "spatial outlyingness function"

$$O_S(\mathbf{x}, F) = \|\mathbf{E}\mathbf{S}(\mathbf{x} - \mathbf{X})\|, \quad \mathbf{x} \in \mathbb{R}^d$$

- ▶ It is (only) orthogonally invariant, for $d \geq 2$.
- ▶ In the *univariate* case, this is $O_1(x, F) = |2F(x) - 1|$.

Example: *Mahalanobis Distance Approach*

- ▶ Substitution of *multivariate* location and spread measures $\mathbf{m}(F)$ and $\mathbf{S}(F)$ in the *scaled deviation* $O_1(\cdot, \cdot)$ yields the popular "*Mahalanobis distance outlyingness*"

$$O_{\text{MD}}(\mathbf{x}, F) = \|\mathbf{S}(F)^{-1/2}(\mathbf{x} - \mathbf{m}(F))\|, \mathbf{x} \in \mathbb{R}^d$$

- ▶ It is *affine invariant*.
- ▶ However, the outlyingness contours are *ellipsoidal* even when the shape of F is not.

Depth Functions: *Order Statistics* in \mathbb{R}^d ?

Depth function: a center-outward ordering $D(\mathbf{x}, F)$ with higher values representing greater “centrality”.

Tukey '75 – Liu '88 – Donoho & Gasko '92 – Vardi & Zhang '00 – Zuo '03

- ▶ Compensates for lack of linear order in \mathbb{R}^d , $d \geq 2$, by *orienting to a “center”*.
- ▶ Maximum depth points define a notion of “center” and a notion of “multidimensional median” \mathbf{M}_F .
- ▶ Can order data points by their sample depths, and define analogues of univariate statistics involving order statistics.

Example & Nonexample

- ▶ *Halfspace Depth* [Tukey, 1975]

$$D_H(\mathbf{x}, F) = \inf\{P(H) : \mathbf{x} \in H \text{ closed halfspace}\}, \mathbf{x} \in \mathbb{R}^d$$

- ▶ However, the *density function* is *not* a depth:
 - ▶ It does not in general measure centrality or outlyingness.
 - ▶ Its interpretation is *local* and has *no global perspective*.
 - ▶ The point of maximality is not interpretable as a center.
 - ▶ For F uniform on $[0, 1]^d$, for example, the density function yields no contours at all.

Quantile Functions: “Quantiles” in \mathbb{R}^d ?

Quantile function $\mathbf{Q}(\mathbf{u}, F)$: attaches to each \mathbf{x} a “quantile representation” indexed by \mathbf{u} in $\mathbb{B}^{d-1}(\mathbf{0})$ with *nested* contours

$$\{\mathbf{Q}(\mathbf{u}, F) : \|\mathbf{u}\| = c\}, 0 \leq c < 1.$$

- ▶ For *quantile-based* inference in \mathbb{R}^d , the “center” $\mathbf{Q}(\mathbf{0}, F)$ should be interpretable as a *d-dimensional median* \mathbf{M}_F .
- ▶ For $\mathbf{u} \neq \mathbf{0}$, the index \mathbf{u} represents *direction* in some sense: e.g., *direction* to $\mathbf{x} = \mathbf{Q}(\mathbf{u}, F)$ from \mathbf{M}_F , or *expected direction* to $\mathbf{x} = \mathbf{Q}(\mathbf{u}, F)$ from random $\mathbf{X} \sim F$.
- ▶ The *magnitude* $\|\mathbf{u}\|$ represents an *outlyingness parameter*.

Example: *Spatial Quantile Function*

- ▶ The spatial quantile function $\mathbf{Q}_S(\mathbf{u}, F)$ gives $\boldsymbol{\theta}$ in \mathbb{R}^d minimizing $E\{\Phi(\mathbf{u}, \mathbf{X} - \boldsymbol{\theta})\}$, where $\Phi(\mathbf{u}, \mathbf{t}) = \|\mathbf{t}\| + \mathbf{u}'\mathbf{t}$
[Dudley and Koltchinskii, 1992, Chaudhuri, 1996, and Koltchinskii, 1997].

- ▶ $\mathbf{Q}_S(\mathbf{u}, F)$ is the solution \mathbf{x} of the equation

$$\mathbf{u} = E\mathbf{S}(\mathbf{x} - \mathbf{X}).$$

- ▶ It is (only) orthogonally equivariant, for $d \geq 2$.
- ▶ $\mathbf{Q}_S(\mathbf{0}, F)$ is the well-known spatial median.
- ▶ For $\mathbf{x} = \mathbf{Q}_S(\mathbf{u}, F)$, we have

$$\|\mathbf{u}\| = \|E\mathbf{S}(\mathbf{x} - \mathbf{X})\| = O_S(\mathbf{x}, F).$$

Centered Rank Function: “Signs”, “Ranks” in \mathbb{R}^d ?

Centered rank function $\mathbf{R}(\mathbf{x}, F)$: takes values in $\mathbb{B}^{d-1}(\mathbf{0})$, with origin $\mathbf{0}$ assigned to a multivariate median $\mathbf{x} = \mathbf{M}_F$, and for other \mathbf{x} denotes a “directional rank” in $\mathbb{B}^{d-1}(\mathbf{0})$.

- ▶ The magnitude $\|\mathbf{R}(\mathbf{x}, F)\|$ measures outlyingness of \mathbf{x} .
- ▶ Univariate case. $R(x, F) = 2F(x) - 1$, with its *sign* giving “direction” (from median $F^{-1}(1/2)$), and its *magnitude* providing the “rank” of x .
- ▶ For testing $H_0 : \mathbf{M}_F = \boldsymbol{\theta}_0$, the *sample version* of $\mathbf{R}(\boldsymbol{\theta}_0, F)$ provides a multivariate version of the *univariate sign test*.

Example: *Spatial Centered Rank Function*

- ▶ The *spatial centered rank function* [Möttönen and Oja, 1995] is

$$\mathbf{R}_S(\mathbf{x}, F) = E\mathbf{S}(\mathbf{x} - \mathbf{X}), \quad \mathbf{x} \in \mathbb{R}^d.$$

- ▶ For testing H_0 : “spatial median = θ_0 ”, the statistic

$$\sum_{i=1}^n \mathbf{S}(\theta_0 - \mathbf{x}_i)$$

provides a *spatial sign test statistic*.

- ▶ This test is (only) *orthogonally invariant*, for $d \geq 2$

Equivalence: the *D-O-Q-R Paradigm*

Depth, outlyingness, quantiles, and ranks in \mathbb{R}^d are equivalent.

- ▶ $D(\mathbf{x}, F)$ and $O(\mathbf{x}, F)$ are equivalent (inversely).
- ▶ $\mathbf{Q}(\mathbf{u}, F)$ and $\mathbf{R}(\mathbf{x}, F)$ are equivalent (inversely).
- ▶ These couplets are linked by
 - a) $O(\mathbf{x}, F) = \|\mathbf{R}(\mathbf{x}, F)\| (= \|\mathbf{u}\|)$,
 - b) $D(\mathbf{x}, F)$ induces a corresponding $\mathbf{Q}(\mathbf{u}, F)$.

Each of D , O , \mathbf{Q} , and \mathbf{R} can generate the others, although they are very different in conceptual meaning and appeal.

Example: *Depth-Induced Quantile Functions*

For $D(\mathbf{x}, F)$ having nested contours enclosing “median” \mathbf{M}_F and bounding “central regions” $\{\mathbf{x} : D(\mathbf{x}, F) \geq \alpha\}$, $\alpha > 0$, the *depth contours induce a quantile representation* for $\mathbf{x} \in \mathbb{R}^d$:

- ▶ For $\mathbf{x} = \mathbf{M}_F$, denote it by $\mathbf{Q}(\mathbf{0}, F)$.
- ▶ For $\mathbf{x} \neq \mathbf{M}_F$, denote it by $\mathbf{Q}(\mathbf{u}, F)$ with $\mathbf{u} = p\mathbf{w}$, where p is the probability weight of the central region with \mathbf{x} on its boundary and \mathbf{w} is the unit vector toward \mathbf{x} from \mathbf{M}_F .

Then $\mathbf{u} = \mathbf{R}(\mathbf{x}, F)$ is *direction toward* $\mathbf{x} = \mathbf{Q}(\mathbf{u}, F)$ from \mathbf{M}_F , and $\|\mathbf{u}\| = \|\mathbf{R}(\mathbf{x}, F)\|$ is the *probability weight of the central region with* $\mathbf{x} = \mathbf{Q}(\mathbf{u}, F)$ *on its boundary*.

Desired *Equivariance* and *Invariance* Properties

- ▶ *How should estimators and test statistics, or depth, outlyingness, quantile, and rank functions, change when the data are transformed to other coordinates?*
- ▶ Quantile functions on \mathbb{R}^d are desirably equivariant, and depth and outlyingness functions should be invariant.
- ▶ The *new quantile representation* of a point \mathbf{x} should be given by the *same transformation* of the original quantile representation, subject to a possible reindexing.
- ▶ The *outlyingness* of \mathbf{x} should remain unchanged.

A Further Point, and *Our Goal*

- ▶ Equivariance/invariance is also technically convenient:
 - ▶ For purposes of relative efficiency comparisons or for setting outlyingness thresholds, for example, one may without loss of generality represent a parametric family by a single member for all numerical computations.
- ▶ *Unqualified insistence* on equivariance/invariance as a principle is *not justified*, however, for it may lead to undue compromises of computational efficiency or robustness.
- ▶ We will formulate equivariance/invariance technically and see how to produce it through suitable standardization.

Equivariance of Multivariate Quantile Functions

Definition. An \mathbb{R}^d -valued quantile function $\mathbf{Q}(\mathbf{u}, F)$, $\mathbf{u} \in \mathbb{B}^{d-1}(\mathbf{0})$, is affine equivariant if, for $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ with any nonsingular $d \times d$ \mathbf{A} and any d -vector \mathbf{b} ,

$$\mathbf{Q}(\mathbf{v}, F_{\mathbf{Y}}) = \mathbf{A}\mathbf{Q}(\mathbf{u}, F_{\mathbf{X}}) + \mathbf{b}, \quad \mathbf{u} \in \mathbb{B}^{d-1}(\mathbf{0})$$

with a $\mathbb{B}^{d-1}(\mathbf{0})$ -valued *re-indexing* $\mathbf{v} = \mathbf{v}(\mathbf{u}, \mathbf{A}, \mathbf{b}, F_{\mathbf{X}})$ which satisfies outlyingness invariance

$$\|\mathbf{v}(\mathbf{u}, \mathbf{A}, \mathbf{b}, F_{\mathbf{X}})\| = \|\mathbf{u}\|, \quad \mathbf{u} \in \mathbb{B}^{d-1}(\mathbf{0})$$

Equivariance of Multivariate Median

For the *median* $\mathbf{Q}(\mathbf{0}, F_{\mathbf{X}})$, the equivariance property may be stated simply

$$\mathbf{Q}(\mathbf{0}, F_{\mathbf{Y}}) = \mathbf{A} \mathbf{Q}(\mathbf{0}, F_{\mathbf{X}}) + \mathbf{b}.$$

Equivariance of Contours

- ▶ Denote the contours of a quantile function $\mathbf{Q}(\cdot, F)$ by

$$\tilde{\mathbf{Q}}(c, F) = \{\mathbf{Q}(\mathbf{u}, F) : \|\mathbf{u}\| = c\}, \quad 0 < c < 1.$$

- ▶ If $\mathbf{Q}(\cdot, F)$ is affine equivariant, then equivalently so are the contours: for $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$,

$$\tilde{\mathbf{Q}}(c, F_{\mathbf{Y}}) = \mathbf{A} \tilde{\mathbf{Q}}(c, F_{\mathbf{X}}) + \mathbf{b}, \quad 0 < c < 1.$$

- ▶ Here the mapping $\mathbf{u} \mapsto \mathbf{v}(\mathbf{u}, \mathbf{A}, \mathbf{b}, F_{\mathbf{X}})$ is implicit.

Equivariance/Invariance of Related Functions

- ▶ Equivariance of $\mathbf{Q}(\cdot, F)$ yields equivariance and invariance properties for the related D, O, and R functions.
- ▶ The definition of the centered rank function as the *inverse of the quantile function* immediately yields *equivariance of $\mathbf{R}(\cdot, F)$* , in the following sense:

$$\mathbf{R}(\mathbf{y}, F_Y) = \mathbf{v}(\mathbf{R}(\mathbf{x}, F_X), \mathbf{A}, \mathbf{b}, F_X)$$

- ▶ In turn, the relation $O(\mathbf{x}, F) = \|\mathbf{R}(\mathbf{x}, F)\|$ yields *invariance of $O(\mathbf{x}, F_X)$* and likewise of $D(\mathbf{x}, F_X)$. (These also follow from the definition of equivariance of $\mathbf{Q}(\cdot, F)$.)

$\mathbf{v}(\mathbf{u}, \mathbf{A}, \mathbf{b}, F_X)$ for *Depth-Induced* $\mathbf{Q}(\mathbf{u}, F)$

- ▶ For an *affine equivariant* depth-induced quantile function $\mathbf{Q}(\cdot, F)$, it follows that $\mathbf{M}_Y = \mathbf{A} \mathbf{M}_X + \mathbf{b}$ and then the unnormalized direction vector toward $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$ from \mathbf{M}_Y is $\mathbf{A}(\mathbf{x} - \mathbf{M}_X)$.
- ▶ Then $\mathbf{R}(\mathbf{y}, F_Y) = c_0 \mathbf{A} \mathbf{R}(\mathbf{x}, F_X)$ for some constant c_0 , or equivalently $\mathbf{v}(\mathbf{u}, \mathbf{A}, \mathbf{b}, F_X) = c_0 \mathbf{A} \mathbf{u}$, where $\mathbf{x} = \mathbf{Q}(\mathbf{u}, F)$.
- ▶ *Outlyingness invariance* then requires $|c_0| = \|\mathbf{u}\| / \|\mathbf{A} \mathbf{u}\|$, yielding, for either choice of sign,

$$\mathbf{v}(\mathbf{u}, \mathbf{A}, \mathbf{b}, F_X) = \pm \frac{\|\mathbf{u}\|}{\|\mathbf{A} \mathbf{u}\|} \mathbf{A} \mathbf{u}.$$

$\mathbf{v}(\mathbf{u}, \mathbf{A}, \mathbf{b}, F_{\mathbf{x}})$ for the *Spatial* $\mathbf{Q}(\mathbf{u}, F)$

- ▶ A well-known limitation of the spatial quantile function is its *orthogonal*, rather than full affine, equivariance.
- ▶ That is, the desired equivariance holds only for \mathbf{A} orthogonal, in which case $\mathbf{v}(\mathbf{u}, \mathbf{A}, \mathbf{b}, F_{\mathbf{x}}) = \pm \mathbf{A}\mathbf{u}$.
- ▶ A point \mathbf{x} labeled a (spatial) “outlier” or “nonoutlier” would have the same classification after *orthogonal* transformation to a new coordinate system but not necessarily after transformation by *heterogeneous scale changes*.

Weak Covariance (WC) Functionals

- ▶ **Definition.** A matrix-valued functional $\mathbf{C}(F)$ is a weak covariance (WC) functional if, for $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ with any nonsingular \mathbf{A} and any \mathbf{b} ,

$$\mathbf{C}(F_{\mathbf{Y}}) = k_1 \mathbf{A} \mathbf{C}(F_{\mathbf{X}}) \mathbf{A}'$$

with $k_1 = k_1(\mathbf{A}, \mathbf{b}, F_{\mathbf{X}})$ a positive scalar function.

- ▶ $k_1(\mathbf{A}, \mathbf{b}, F_{\mathbf{X}}) = 1$ gives the usual “*covariance functional*” .
[e.g., Lopushaä and Rousseeuw, 1991]
- ▶ A WC functional is also known as a “*shape functional*” .
[Paindaveine, 2008; Tyler, Critchley, Dümbgen, and Oja, 2009]

Making Spatial Quantiles *Affine* Equivariant

- ▶ **Theorem.** For any WC functional $\mathbf{C}(F)$, the quantile function

$$\mathbf{Q}_{MS}(\mathbf{u}, F_{\mathbf{X}}) = \mathbf{C}(F_{\mathbf{X}})^{1/2} \mathbf{Q}_S(\mathbf{u}, F_{\mathbf{C}(F_{\mathbf{X}})^{-1/2}\mathbf{X}})$$

is *affine equivariant*.

- ▶ The relevant re-indexing for $\mathbf{X} \mapsto \mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ is

$$\mathbf{v}(\mathbf{u}, \mathbf{A}, \mathbf{b}, F_{\mathbf{X}}) = \tilde{\mathbf{A}}(\mathbf{A}, F_{\mathbf{X}}) \mathbf{u},$$

where $\tilde{\mathbf{A}}(\mathbf{A}, F_{\mathbf{X}}) = (\mathbf{A} \mathbf{C}(F_{\mathbf{X}}) \mathbf{A}')^{1/2} (\mathbf{A}')^{-1} \mathbf{C}(F_{\mathbf{X}})^{-1/2}$,
which is *orthogonal*.

Comments on the Quantile Function $Q_{MS}(\cdot, F)$

- ▶ In homage to Mahalanobis, who promoted the role of *standardization* in multivariate analysis, we may call $Q_{MS}(\cdot, F)$ the “Mahalanobis spatial quantile function” corresponding to $\mathbf{C}(\cdot)$ -standardization.
- ▶ The favorable properties of the spatial quantile function $Q_S(\cdot, F)$ carry over to $Q_{MS}(\cdot, F)$, subject to the quality of choice of the WC functional $\mathbf{C}(F)$.
- ▶ The contours of $Q_{MS}(\cdot, F)$ *need not be elliptical*, in contrast to the “Mahalanobis distance quantile function” $Q_{MD}(\cdot, F)$ corresponding to $O_{MD}(\cdot, F)$ discussed earlier.

Transformation-Retransformation (TR) Functionals

- ▶ **Definition.** A matrix-valued functional $\mathbf{M}(F)$ is a transformation-retransformation (TR) functional if, for $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ with any nonsingular \mathbf{A} and any \mathbf{b} ,

$$\mathbf{A}'\mathbf{M}(F_{\mathbf{Y}})'\mathbf{M}(F_{\mathbf{Y}})\mathbf{A} = k_2 \mathbf{M}(F_{\mathbf{X}})'\mathbf{M}(F_{\mathbf{X}})$$

with $k_2 = k_2(\mathbf{A}, \mathbf{b}, F_{\mathbf{X}})$ a positive scalar function.

[Chakraborty and Chaudhuri, 1996; Randles, 2000]

- ▶ TR approaches modify *estimation (testing)* procedures to achieve (*hopefully*) full *affine equivariance (invariance)*.
 - ▶ Carry out the procedure on transformed data $\mathbf{M}(\mathbb{X}_n)\mathbb{X}_n$, then retransform to original coordinates via $\mathbf{M}(\mathbb{X}_n)^{-1}$.
 - ▶ Verify that the equivariance (invariance) holds.

Connection between TR and WC Functionals

- ▶ **Theorem.** *Every TR functional $\mathbf{M}(F)$ is equivalent to a WC functional, and conversely.*
 - ▶ *For a TR $\mathbf{M}(F)$, $\mathbf{C}(F) = (\mathbf{M}(F)' \mathbf{M}(F))^{-1}$ is a WC fcnl.*
 - ▶ *For a WC $\mathbf{C}(F)$, $\mathbf{M}(F) = \mathbf{C}(F)^{-1/2}$ is a TR fcnl.*
- ▶ Selection of a TR functional is merely an indirect but equivalent way to select a WC functional.
- ▶ Extensive literature on covariance functionals provides many choices meeting various criteria of robustness and computational efficiency.

TR Connection between $\mathbf{Q}_{MS}(\cdot, F)$ and $\mathbf{Q}_S(\cdot, F)$

- ▶ The defining formula

$$\mathbf{Q}_{MS}(\mathbf{u}, F_{\mathbf{X}}) = \mathbf{C}(F_{\mathbf{X}})^{1/2} \mathbf{Q}_S(\mathbf{u}, F_{\mathbf{C}(F_{\mathbf{X}})^{-1/2} \mathbf{X}})$$

exhibits a “*transformation-retransformation*” (TR) representation of $\mathbf{Q}_{MS}(\cdot, F)$ in terms of $\mathbf{Q}_S(\cdot, F)$.

- ▶ The outlyingness function corresponding to $\mathbf{Q}_{MS}(\cdot, F)$ via the D-O-Q-R paradigm is then given by

$$O_{MS}(\mathbf{x}, F_{\mathbf{X}}) = O_S\left(\mathbf{C}(F_{\mathbf{X}})^{-1/2} \mathbf{x}, F_{\mathbf{C}(F_{\mathbf{X}})^{-1/2} \mathbf{X}}\right)$$

and is *affine invariant*.

TR *Sample Version* of $\mathbf{Q}_S(\cdot, F)$

- ▶ Advocates of $\mathbf{Q}_S(\cdot, F)$ reject the *idea* of $\mathbf{Q}_{MS}(\cdot, F)$.
- ▶ *In practice*, however, instead of the sample version $\mathbf{Q}_S(\cdot, \mathbb{X}_N)$ they use, for a particular TR fcnl $\mathbf{M}_0(\mathbb{X}_n)$,

$$\mathbf{Q}^{(\text{TR})}(\mathbf{u}, \mathbb{X}_n) = \mathbf{M}_0(\mathbb{X}_n)^{-1} \mathbf{Q}_S(\mathbf{u}, \mathbf{M}_0(\mathbb{X}_n) \mathbb{X}_n),$$

which achieves *full affine equivariance*.

[see Chakraborty, Chaudhuri, and Oja, 1998, and Chakraborty, 2001]

- ▶ What $\mathbf{Q}^{(\text{TR})}(\mathbf{u}, \mathbb{X}_n)$ actually *estimates*, however, is not $\mathbf{Q}_S(\mathbf{u}, F)$, but rather $\mathbf{Q}_{MS}(\mathbf{u}, F)$ with $\mathbf{C}(F) = \mathbf{M}_0(F)^{-1/2}$.

Invariant Coordinate System (ICS) Functionals

Definition. A matrix-valued functional $\mathbf{D}(F)$ is an invariant coordinate system (ICS) functional if the $\mathbf{D}(\cdot)$ -standardization of \mathbf{X}

$$\mathbf{D}(F_{\mathbf{X}})\mathbf{X}$$

remains unaltered after affine transformation to $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ followed by $\mathbf{D}(\cdot)$ -standardization of \mathbf{Y} to

$$\mathbf{D}(F_{\mathbf{Y}})\mathbf{Y}$$

except for coordinatewise scale changes, sign changes and translations.

[Tyler, Critchley, Dümbgen, and Oja, 2009]

Practical Interpretation of ICS-Standardization

- ▶ With $\mathbf{D}(\cdot)$ an ICS functional, any *geometric structures or patterns* identified in a $\mathbf{D}(\cdot)$ -standardized data set

$$\mathbf{D}(\mathbb{X}_n)\mathbb{X}_n$$

remain unaltered after *affine transformation* to $\mathbb{Y}_n = \mathbf{A}\mathbb{X}_n + \mathbf{b}$ followed by $\mathbf{D}(\cdot)$ -standardization to

$$\mathbf{D}(\mathbb{Y}_n)\mathbb{Y}_n$$

except for *coordinatewise scale changes, sign changes and translations*.

- ▶ Some applications, however, for example *outlyingness*, call for homogeneity of scale changes and sign changes.

Construction of ICS Functionals

Tyler, Critchley, Dümbgen, and Oja (2009) provide an approach for construction of ICS functionals.

- ▶ Let $\mathbf{V}_1(F)$ and $\mathbf{V}_2(F)$ be two WC functionals with the eigenvalues of $\mathbf{V}_1(F)^{-1}\mathbf{V}_2(F)$ all distinct. Then the matrix of corresponding eigenvectors is an ICS functional.
- ▶ Extension for multiplicities among eigenvalues is given.
- ▶ Various choices of $\mathbf{V}_1(F)$ and $\mathbf{V}_2(F)$ are considered.

Strong ICS Functionals

Definition. An ICS functional $\mathbf{D}(F)$ has *Structure A* if, for $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ with any nonsingular \mathbf{A} and any \mathbf{b} ,

$$\mathbf{D}(F_{\mathbf{Y}}) = k_3 \mathbf{J} \mathbf{D}(F_{\mathbf{X}}) \mathbf{A}^{-1}$$

with $k_3 = k_3(\mathbf{A}, \mathbf{b}, F_{\mathbf{X}})$ a positive scalar function and $\mathbf{J} = \mathbf{J}(\mathbf{A}, \mathbf{b}, F_{\mathbf{X}})$ a *sign change matrix* (diagonal with ± 1).

Definition. A strong ICS functional is a Structure A ICS functional with $\mathbf{J} = \mathbf{I}_d$.

For a *strong* ICS functional, both scale changes and sign changes are *homogeneous*.

Connection between ICS and TR Functionals

Theorem. *Every ICS functional $\mathbf{D}(F)$ with Structure A is a TR functional (and thus $(\mathbf{D}(F)'\mathbf{D}(F))^{-1}$ is a WC functional).*

Examples of Strong ICS Functionals

- ▶ The TR functional \mathbf{M}_0 mentioned earlier with $\mathbf{Q}_S(\cdot, \mathbb{X}_N)$ is a strong ICS functional with $\mathbf{J} = \mathbf{I}_d$. It is constructed by a method quite different from the above.
- ▶ The well-known and oft-used Tyler (1987) scatter functional is also a TR functional but not a strong ICS functional.
- ▶ A symmetrized version of the Tyler functional given by Dümbgen (1998) does not involve a location functional and also is a TR functional.

Key Property of Strong ICS Functionals

- ▶ A *strong* ICS functional $\mathbf{D}(F)$ satisfies, for $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$,

$$\mathbf{D}(F_{\mathbf{Y}})\mathbf{Y} = k_3 \mathbf{D}(F_{\mathbf{X}})\mathbf{X} + \mathbf{c}$$

with $\mathbf{c} = \mathbf{c}(\mathbf{A}, \mathbf{b}, F_{\mathbf{X}}) = k_3 \mathbf{D}(F_{\mathbf{X}})\mathbf{A}^{-1}\mathbf{b}$, a *constant*.

- ▶ Thus the new $\mathbf{D}(\cdot)$ -standardized coordinates $\mathbf{D}(F_{\mathbf{Y}})\mathbf{Y}$ agree with the original $\mathbf{D}(\cdot)$ -standardized coordinates $\mathbf{D}(F_{\mathbf{X}})\mathbf{X}$, except for a *homogeneous scale change* and a *translation*.

Practical Interpretation in the *Strong* ICS Case

With $\mathbf{D}(\cdot)$ a *strong* ICS functional,

$$\mathbf{D}(\mathbb{X}_n)\mathbb{X}_n$$

remains unaltered after *affine transformation* to $\mathbb{Y}_n = \mathbf{A}\mathbb{X}_n + \mathbf{b}$ followed by $\mathbf{D}(\cdot)$ -*standardization* to

$$\mathbf{D}(\mathbb{Y}_n)\mathbb{Y}_n,$$

except for a *homogeneous scale change* and a *translation*.

Affine Invariance via Strong ICS Functionals

Theorem. *Let $T(\mathbf{x}, F)$ be a real-valued functional of \mathbf{x} and F that is invariant under homogeneous scale change and translation of \mathbf{x} , in the sense that*

$$T(c\mathbf{x} + \mathbf{b}, F_{c\mathbf{x} + \mathbf{b}}) = T(\mathbf{x}, F_{\mathbf{x}})$$

for any scalar c and any vector \mathbf{b} . Let $\mathbf{D}(F)$ be a strong ICS functional. Then the functional

$$T(\mathbf{D}(F_{\mathbf{x}})\mathbf{x}, F_{\mathbf{D}(F_{\mathbf{x}})\mathbf{x}})$$

is affine invariant.

Example Neding the Theorem

Scaled-deviation outlyingness for a single projection \mathbf{u}_0

$$T(\mathbf{x}, F_{\mathbf{x}}) = \left| \frac{\mathbf{u}'_0 \mathbf{x} - \mu(F_{\mathbf{u}'_0 \mathbf{x}})}{\sigma(F_{\mathbf{u}'_0 \mathbf{x}})} \right|$$

With Multivariate Data, First Standardize!

- ▶ *Example.* To make the *spatial quantile function* affine equivariant, first standardize with a *TR functional*.
 - ▶ Choosing a *strong ICS functional* yields added benefits.
- ▶ *Example.* In *projection pursuit* methods with multivariate data, univariate standardization after projection does not in general produce desired affine invariance. Standardize first, using a *strong ICS functional*.
 - ▶ E.g., to formulate *outlyingness* of \mathbf{x} in \mathbb{R}^d by a quadratic form based on univariate scaled deviations of projections of \mathbf{x} on just finitely many directions $\{\mathbf{u}_1, \dots, \mathbf{u}_s\}$, use $\mathbf{u}'_i \mathbf{D}(F_{\mathbf{X}}) \mathbf{x}$ instead of $\mathbf{u}'_i \mathbf{x}$, $1 \leq i \leq s$, etc.

Many Types of Depth Function, for Example

- ▶ *Spatial depth* (Dudley and Koltchinski, 1992; Möttönen and Oja, 1995; Chaudhuri, 1996; Vardi and Zhang, 2000)
- ▶ *Simplicial depth* (Liu, 1988)
- ▶ *Majority depth* (Singh, 1991)
- ▶ *Projection depth* (Liu, 1992; Zuo, 2003)
- ▶ *Simplicial volume depth* (based on Oja, 1983)
- ▶ L^p *depth* (Zuo and Serfling, 2000)
- ▶ *Mahalanobis depth* (Liu and Singh, 1993)
- ▶ *Zonoid depth* (Koshevoy and Mosler, 1997)
- ▶ And more ... a growing industry!

Desirable Properties of Depth Functions

- ▶ *Affine invariance.* $D(\mathbf{x}, F)$ independent of coordinate system.
- ▶ *Maximality at center.* $D(\mathbf{x}, F)$ maximal at a “center”.
- ▶ *Symmetry.* If F symmetric about θ , so is $D(\mathbf{x}, F)$.
- ▶ *Decreasing along rays.* $D(\mathbf{x}, F)$ decreases along each ray from deepest point.
- ▶ *Vanishing at infinity.* $D(\mathbf{x}, P) \rightarrow 0$ as $\|\mathbf{x}\| \rightarrow \infty$.
- ▶ *Continuity of $D(\mathbf{x}, P)$ as a function of \mathbf{x} .* Or merely upper semicontinuity.
- ▶ *Continuity of $D(\mathbf{x}, P)$ as a functional of P .*
- ▶ *Quasi-concavity as a function of \mathbf{x} .* $\{\mathbf{x}: D(\mathbf{x}, P) \geq c\}$ is convex for each real c .

What We Can Compute from a Depth Function

- ▶ *Contours.*
- ▶ *Depth-order statistics:* $\mathbf{X}_{[1]}, \dots, \mathbf{X}_{[n]}$ (center-outward)
- ▶ *Depth-weighted location functionals:*

$$\frac{\int_{\mathbb{R}^d} \mathbf{x} W(D(\mathbf{x}, F)) dF(\mathbf{x})}{\int_{\mathbb{R}^d} W(D(\mathbf{x}, F)) dF(\mathbf{x})}.$$

- ▶ *Scale curves:* Volume within contour vs probability weight.
- ▶ *Skewness functionals.* Scaled difference of two location functionals.
- ▶ *Kurtosis functionals.* Via transformation of scale curve.
- ▶ *And more ...*

Depth-Based Statistical Procedures

- ▶ *Bagplots, sunburst plots.* Extend boxplot to $d = 2$ and 3, using contours and rays to outlying points. Answers the question: *where is the “middle half” of the data?*
- ▶ *DD, PP, QQ plots.* Compare two samples by a plot of depth values of combined sample, or of the volumes of sample central regions versus each other, or of kurtosis curves against each other, or of depth-based quantiles versus each other.
- ▶ *Comparison of several distributions.* Plot scale curves in a single exhibit. Or kurtosis curves.
- ▶ *Nonparametric description of multivariate distributions.* Measures of location, spread, asymmetry, kurtosis.

And More ...

- ▶ *Testing multivariate symmetry.* For *spherical symmetry*, plot fraction of data in smallest sphere containing the p th sample central region. For *central symmetry*, plot fraction of data within the intersection of p th sample central region and its reflection.
- ▶ *Diagnosis of nonnormality.* Use trimmed depth-weighted scatter matrix. Or kurtosis curve.
- ▶ *Outlier identification.*
- ▶ *P-values via bootstrap and data depth.*
- ▶ *One- and multi-sample multivariate rank statistics defined on depth-based ranks.*

And More ...

- ▶ *Statistical process control procedures.* Can use depth-based ranks for monitoring and thresholding with multivariate data using univariate quality control procedures.

Compare outlyingness of a new observation relative to in-control reference point cloud.
- ▶ *Multivariate density estimation by probing depth.*
- ▶ *Depth-based quality indices.*
- ▶ *Depth-based classification rules.*
- ▶ *Depth-based cluster analysis.*

Contexts of Application

Depth-based methods provide new competitors for standard approaches having diverse applications:

- ▶ *Exploratory data analysis*
- ▶ *Multi-sample inference*
- ▶ *Regression.*
- ▶ *Classification, clustering, discrimination*
- ▶ *Directional analysis*
- ▶ *Multivariate density estimation*

Contexts of Application

Further, in some application contexts, depth-based methods are especially natural or advantageous:

- ▶ *Monitoring of aviation safety data.*
- ▶ *Industrial quality control.*
- ▶ *Measuring economic disparity and concentration*
- ▶ *Social choice and voting.* Ordering social choices by voters' preferences, and finding "median voter".
- ▶ *Game-theoretic analysis of competition.*

Computational Burden Involves ... Computational Geometry

- ▶ *Halfspace and simplicial depth: $O(n^{d-1} \log n)$.*
- ▶ *Nested convex hulls of fractions of data: $O(n^{d-1} \log n)$.*
- ▶ *Bivariate halfspace contours, bagplot, & depths: $O(n^2)$.*
- ▶ *Deepest regression hyperplane: $O(n^d)$.*
- ▶ *Regression depth of a k -flat: $O(n^{d-2} + n \log n)$, for $1 \leq k \leq d - 2$.*
- ▶ *Stahel-Donoho estimators: $O(n^{d+1})$.*
- ▶ *Volume of sample p th central region: nontrivial.*
- ▶ *Simulation and bootstrap: many samples needed.*
- ▶ *Issues of NP-hard, NP-complete, coNP-complete, etc.*

Some Desirable Results for Sample Versions

- ▶ *Consistency:*

$$\|D(\mathbf{x}, \hat{F}_n) - D(\mathbf{x}, F)\|_\infty \rightarrow 0, \quad n \rightarrow \infty.$$

- ▶ *Weak convergence:*

$$n^{1/2}[D(\mathbf{x}, \hat{F}_n) - D(\mathbf{x}, F)] \text{ converges weakly.}$$

- ▶ *Convergence of sample central regions.*
- ▶ *Convergence of functionals of sample depth or of sample central regions.*
- ▶ *Robustness of functionals of sample depth or of sample central regions.* E.g., breakdown points, influence curves.
- ▶ *Favorable behavior of bootstrap.*

Consistency of Depth Process

We want

$$\|D(\mathbf{x}, \hat{F}_n) - D(\mathbf{x}, F)\|_\infty \rightarrow 0, \quad n \rightarrow \infty.$$

- ▶ For *halfspace depth*: Donoho and Gasko (1992).
- ▶ For *simplicial depth*: Liu (1990), Dümbgen (1990), and Arcones and Giné (1993).
- ▶ For *majority and Mahalanobis depths*: Liu and Singh (1993). For *projection depth and Type D depths*: Zuo and Serfling (2000).

Weak Convergence of Depth Process

We want

$$n^{1/2}[D(\mathbf{x}, \hat{F}_n) - D(\mathbf{x}, F)] \text{ converges weakly.}$$

- ▶ For *halfspace depth*: Massé (2004)
- ▶ For *simplicial depth*: Dümbgen (1990), and Arcones and Giné (1993). Note: This depth process is a special case of *U-process*.
- ▶

Convergence of Sample Central Regions

- ▶ We desire convergence of contours in *Hausdorff distance*.
- ▶ For *halfspace depth*: Eddy (1985), Nolan (1992), Donoho and Gasko (1992), Massé and Theodorescu (1994), and He and Wang (1997).
- ▶ For *simplicial, projection, and Type D depths*: Zuo and Serfling (2000).

Asymptotic Normality of Location Estimator $L(\hat{F}_n)$

$L(F) =$

$$\int_{\mathbb{R}^d} \mathbf{x} W(D(\mathbf{x}, F)) dF(\mathbf{x}) \Big/ \int_{\mathbb{R}^d} W(D(\mathbf{x}, F)) dF(\mathbf{x}) .$$

- ▶ For *simplicial depth*: Dümbgen (1992).
- ▶ For *halfspace depth*: Massé (1999, 2004).
- ▶ For *generalized Tukey depth*: Zhang (2002).
- ▶ For *projection depth, Mahalanobis distance depth, and others*: Zuo, Cui, and He (2004).

Asymptotic Normality of Scatter Estimator $S(\hat{F}_n)$

$S(F) =$

$$\int_{\mathbb{R}^d} (\mathbf{x} - L(F))(\mathbf{x} - L(F))' w(D(\mathbf{x}, F)) dF(\mathbf{x}) \bigg/ \int_{\mathbb{R}^d} w(D(\mathbf{x}, F)) dF(\mathbf{x}) .$$

- ▶ Zuo and Cui (2005).

Weak Convergence of Sample Scale Functions and Similar Constructions

- ▶ The volume functional may be written as a “*generalized (univariate) quantile function*” in the sense of Einmahl and Mason (1992).
- ▶ Accordingly, appropriate sample versions may be treated as generalized quantile processes, but the formulation of sample versions that fit into the Einmahl/Mason framework is problematic.
- ▶ Partial results are available in Serfling (2002) and Wang and Serfling (2006).

Bahadur-Kiefer Representation for Spatial Quantile Function

- ▶ Chaudhuri (1992, 1996) and Koltchinskii (1994).
- ▶ Target: extend to *Mahalanobis spatial quantile function*(s).
- ▶ Zhou and Serfling (2008) get B-K representations for *spatial U-quantiles*, based on \mathbb{R}^d -valued $\mathbf{h}(x_1, \dots, x_m)$.
- ▶ Target: extend to *Mahalanobis spatial U-quantile function*(s).
- ▶ Target: obtain Bahadur-Kiefer representations in some form for other quantile functions.

Bahadur-Kiefer Representations Have Several Applications

- ▶ *Joint asymptotic normality of a vector of sample quantiles*
- ▶ *Linkage between estimators and test statistics.*
 - ▶ For certain *spatial U-quantile functions*, the related test statistics via the Bahadur-Kiefer representation represent *multivariate Wilcoxon signed rank statistics* and have been studied in Möttönen and Oja (1995), Möttönen, Oja, and Tienari (1997), and Möttönen, Oja, and Serfling (2005).
- ▶ *Almost sure convergence and law of iterated logarithm for sample quantile function.*

A Few Comments

- ▶ The preceding examples are representative illustrations, with the emphasis on depth functions for convenience.
- ▶ The “D” examples imply by the D-O-Q-R paradigm examples for O-Q-R.
- ▶ The relevant asymptotics includes the complication of dealing with sample versions of any standardizing functionals.

A Few More Comments

- ▶ Our treatment has confined to *location depth*, i.e., to depth functions such that the maximal depth point is a *location parameter*.
- ▶ An extended treatment would cover depth functions on the *parameter space*, such that the maximal depth point is a *dispersion matrix* or a *regression hyperplane*, for example.

Final Comments

- ▶ Asymptotic problems in the D-O-Q-R landscape are diverse and complex with *open gaps and needed extensions*.
- ▶ The asymptotics of D-O-Q-R methods in nonparametric multivariate analysis (and beyond) is a *challenge!*
- ▶ An *apologia*: This presentation is brief and uneven, with much relevant work unmentioned. It provides a sketchy overview of the *landscape* and suppresses technical detail.
- ▶ In preparation, however: *Depth and Quantile Functions in Nonparametric Multivariate Analysis*, Springer

Acknowledgments

The speaker thanks G. L. Thompson, Marc Hallin, Satyaki Mazumder, Hannu Oja, Davy Paindaveine Ron Randles, and many others including anonymous commentators, for very thoughtful, stimulating, and helpful remarks.

Also, support by NSF grants DMS-0103698 and DMS-0805786 and NSA Grant H98230-08-1-0106 is greatly appreciated.