Groupe de Recherche en
Économie Théorique et Appliquée

Conditional Spatial Quantile: Characterization and Nonparametric Estimation

Mohamed CHAOUCH

IMB, Université de Bourgogne

Ali GANNOUN

CNAM, Paris

Jérôme SARACCO

GREThA UMR CNRS 5113

Cahiers du GREThA
n° 2008-10

GRETHA UMR CNRS 5113
Université Montesquieu Bordeaux IV
Avenue Léon Duguit - 33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75 - Fax : +33 (0)5.56.84.86.47 - www.gretha.fr

Caractérisation de quantiles spatiaux conditionnels et estimation non paramétrique

Résumé

Dans le cadre d'études économiques, biomédicales ou industrielles par exemple, on cherche souvent à déterminer le quantile d'un vecteur aléatoire conditionnellement à un autre. On parle alors de quantiles spatiaux conditionnels. Dans cet article, nous traitons dans un premier temps le cas de quantiles spatiaux, puis celui de quantiles spatiaux conditionnels. Il est à noter que l'absence de relation d'ordre total dans un espace multidimensionnel ne va pas permettre de généraliser directement la notion de quantiles univariés (conditionnels ou non) au cas des quantiles spatiaux ou multivariés. Nous nous focalisons ici sur la notion de quantile spatial telle qu'elle a été proposée par Chaudhuri (1996) et nous donnons les estimateurs correspondants. A cet effet, nous présentons deux algorithmes permettant le calcul des estimateurs proposés. Une implémentation sous le logiciel R de ces algorithmes a été mise en oeuvre. Pour finir, nous illustrons les différentes notions de quantiles spatiaux non conditionnels et conditionnels l'aide de jeux de données simulées.

Mots-clés : Quantile spatial, Quantile spatial conditionnel, Estimateur à noyau, Contours

Conditional Spatial Quantile: Characterization and Nonparametric Estimation

Abstract

Conditional quantiles are required in various economic, biomedical or industrial problems. Lack of objective basis for ordering multivariate observations is a major problem in extending the notion of quantiles or conditional quantiles (also called regression quantiles) in a multidimensional setting. We first recall some characterizations of the unconditional spatial quantiles and the corresponding estimators. Then, we consider the conditional case. In this work, we focus our study on the geometric (or spatial) notion of quantiles introduced by Chaudhuri (1992a, 1996). We generalize, in the conditional framework, the Theorem 2.1.2 of Chaudhuri (1996), and we present algorithms allowing the calculation of the unconditional and conditional spatial quantile estimators. Finally, these various concepts are illustrated using simulated data.

Keywords: Conditional Spatial Quantile, Contours, Kernel Estimators, Spatial Quantile

JEL : C14 ; C63

GRETHA UMR CNRS 5113
Université Montesquieu Bordeaux IV
Avenue Léon Duguit - 33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75 - Fax : +33 (0)5.56.84.86.47 - www.gretha.fr

Conditional Spatial Quantile: Characterization and Nonparametric Estimation

Mohamed Chaouch[1] , Ali Gannoun[2] & Jérôme Saracco[3]

[1] Institut de Mathématiques de Bourgogne (UMR CNRS 5584)
Université de Bourgogne
[2] CEDRIC, Chaire de Statistique Appliquée
CNAM, Paris
[3] GREThA (UMR CNRS 5113)
Université Montesquieu - Bordeaux IV

(e-mail : Mohamed.Chaouch@u-bourgogne.fr ; ali.gannoun@cnam.fr ;
Jerome.Saracco@u-bordeaux4.fr)

Abstract: Conditional quantiles are required in various economic, biomedical or industrial problems. Lack of objective basis for ordering multivariate observations is a major problem in extending the notion of quantiles or conditional quantiles (also called regression quantiles) in a multidimensional setting. We recall in first time some characterizations of the unconditional spatial quantiles and the corresponding estimators. Then, we consider the conditional case. In this work, we focus our study on the geometric (or spatial) notion of quantiles introduced by Chaudhuri (1992a, 1996). We generalize, in the conditional framework, the Theorem 2.1.2 of Chaudhuri (1996), and we present algorithms allowing the calculation of the unconditional and conditional spatial quantile estimators. Finally, these various concepts are illustrated using simulated data.

Keywords: Conditional Spatial Quantile, Contours, Kernel Estimators, Spatial Quantile.

Titre en français : **Caractérisation de quantiles spatiaux conditionnels et estimation non paramétrique**

Résumé : Dans le cadre d'études économiques, biomédicales ou industrielles par exemple, on cherche souvent à déterminer le quantile d'un vecteur aléatoire conditionnellement à un autre. On parle alors de quantiles spatiaux conditionnels. Dans cet article, nous traitons dans un premier temps le cas de quantiles spatiaux, puis celui de quantiles spatiaux conditionnels. Il est à noter que l'absence de relation d'ordre total dans un espace multidimensionnel ne va pas permettre de généraliser directement la notion de quantiles univariés(conditionnels ou non) au cas des quantiles spatiaux ou multivariés. Nous nous focalisons ici sur la notion de quantile spatial telle qu'elle a été proposée par Chaudhuri (1996) et nous donnons les estimateurs correspondants. A cet effet, nous présentons deux algorithmes permettant le calcul des estimateurs proposés. Une implémentation sous le logiciel **R** de ces algorithmes a été mise en oeuvre. Pour finir, nous illustrons les différentes notions de quantiles spatiaux non conditionnels et conditionnels à l'aide de jeux de données simulées.

Mots-clé: Quantile spatial, Quantile spatial conditionnel, Estimateur à noyau, Contours.

1 Introduction

Quantiles of univariate data are frequently used to construct popular descriptive statistics. For example, the median is a robust indicator of the central tendency of a population and the interquartile range is a good one's for its dispersion. In addition, quantiles have been used in regression setup (called *"regression quantiles"*) (see Efron, 1991 and Koenker and Basset, 1978) with a univariate response to get robust estimators of parameters in linear models (see Chaudhuri, 1992b and Koenker and Portnoy, 1987). From a practical point of view, quantiles are computed according to an order criterion. Because this order is not total on $\mathbb{R}^d$, an extension of the classical quantile definition in the case when observations are in $\mathbb{R}^d$ can be only partial. It acts in this case of the quantile vector (called *arithmetic*) whose components are the marginal classical quantiles. This definition suffers from several weaknesses. In particular, it is not invariant by rotation and it does not take account of the possible existence of correlations between the different components of the vectors of observations (see Chakraborty, 2001).

Some authors are interested to the problem of ordering multivariate observations and they have gove several techniques, for example Barnett (1976) Plackett (1976) and Reiss (1989). In statistical literature we find some approaches proposed to define quantiles for multivariate data. For example Eddy (1985) defined mutivariate quantiles using nested sequence of sets and Brown and Hettmansperger (1987, 1989) introduced bivariate quantiles based on the definition of Oja's median (see Oja, 1983). Recently, Donoho and Gasko (1992), Liu, Parelius and Singh (1999) and Zuo and Serfling (2000) defined multivariate quantile using different depth functions and Abdous and Theodorescu (1992), Chaudhuri (1996) and Koltchinskii (1997) defined them with a class of M-estimates (see Serfling, 1980). The definition of multivariate quantile proposed by Chaudhuri (1996) (called *geometric*) is equivariant under any homogeneous scale transformation of the coordinates of the multivariate observations (Chaudhuri, 1996). From now on, we will speak about spatial quantiles to refer to this definition.

Within the biomedical studies framework, a variable of interest $\mathbf{Y}$ with values in $\mathbb{R}^d$ (for example blood pressure with its two components: systolic and diastolic pressures) can be concomitant with an explanatory variable $\mathbf{X}$ with values in $\mathbb{R}^s$ (for example the age and the weight of the patient). In this case, we are brought to seek the conditional spatial quantile of $\mathbf{Y}$ given $\mathbf{X}$.

This paper is organized as follows. In Section 2, we recall characterizations of the univariate quantile function. They are generalized in Section 3 to define the spatial quantile. We present then an algorithm allowing to calculate its estimator. In Section 4, we present the theoretical conditional spatial quantiles and their estimators. A calculation algorithm of these estimators is also exposed. Examples on simulated data are given in Section 5 in order to illustrate the numerical behaviors of the estimators. Finally technical proofs are deferred in the Appendix.

2 Univariate quantiles

2.1 Definition

Let $Y \in \mathbb{R}$ be an univariate random variable, and let F be its cumulative distribution function ($c.d.f.$) The quantile function is defined as the inverse of the $c.d.f.$ When F is a monotonically increasing function, its inverse can be defined without ambiguity, but it remains constant on all intervals on which the random variable does not take values. In a general way, the quantile function of Y is noted $Q_F(.)$ and it is defined, for $p \in (0,1)$, such as:

$$Q_F(p) = F^{-1}(p) = \inf\{y : F(y) \geq p\}. \tag{1}$$

2.2 Two characterizations of univariate quantiles

2.2.1 Characterization by equation root

Let $Q(.)$ be a function defined on the interval $(-1,1)$ as:

$$Q(u) = F^{-1}\left(\frac{1+u}{2}\right).$$

The function $Q(.)$ is named "median-centred quantile function" and it satisfies:

- for $u = 0$, $Q(0)$ is the (classical) median (the quantile of order $p = 1/2$),
- $Q^{-1}(y) = 2F(y) - 1$.

Now let S be the following function: $S(y - Y) = \begin{cases} 1 & \text{if } \ y - Y \geq 0, \\ -1 & \text{if } \ y - Y < 0. \end{cases}$

The quantile $Q_F(p)$ is the root of the equation

$$E(S(y - Y)) - (2p - 1) = 0. \tag{2}$$

Proof

Let $u = 2p - 1$, using the above definition of $Q(.)$, we have:

$$\begin{aligned} F\left(F^{-1}(p)\right) - p &= P\left(Y \leq F^{-1}(p)\right) - p \\ &= E(\mathbb{1}_{\{Y \leq F^{-1}(p)\}}) - p \\ &= E(\mathbb{1}_{\{Y \leq F^{-1}\left(\frac{1+u}{2}\right)\}}) - \tfrac{1+u}{2} \\ &= E(\mathbb{1}_{\{Y \leq Q(u)\}}) - \tfrac{1+u}{2} \\ &= E(\mathbb{1}_{\{Q(u) - Y \geq 0\}} - \tfrac{1+u}{2}) \\ &= \tfrac{1}{2}E([2\mathbb{1}_{\{Q(u) - Y \geq 0\}} - 1] - u) \\ &= \tfrac{1}{2}E(S(Q(u) - Y) - u) \end{aligned} \qquad \square$$

Because $F\left(F^{-1}(p)\right) - p = 0$, we deduce that $Q_F(p) = Q(u)$ is the solution y of the equation (2).

2.2.2 Characterization by minimization approach

Using Ferguson(1967) and Koenker and Basset (1978), the quantile can be defined as the solution of the following minimization problem. Let $p \in (0,1)$ a fixed probability. For $t \in \mathbb{R}$, let $\phi(2p-1,t) = |t| + (2p-1)t$ the so-called *loss* function. The quantile function of Y is noted $Q_M(.)$ and it is defined such that

$$Q_M(p) = \arg\min_{\theta\in\mathbb{R}} E\{\phi(2p-1, Y-\theta)\} = \arg\min_{\theta\in\mathbb{R}} \int_{\mathbb{R}} \left(|y-\theta| + (2p-1)(y-\theta)\right) F(dy). \quad (3)$$

It is easy to check that, for $u = 2p-1$, the quantile $Q_M(p)$ may be also represented as the solution y of the equation $E(S(y-Y)) = u$. That is $Q_M(p) = Q(u)$ with $u = 2p-1$.

2.2.3 Remarks

1) For a fixed p, $Q_F(p) = Q_M(p) = Q(u)$ when $u = 2p-1$.
2) The function $Q^{-1}(.)$ is called "centred rank function". The sign of $u = Q^{-1}(y)$ indicates the position of the point y compared to the median: if u is negative (resp. positive), y is on the left (resp. on the right) of the median. Moreover, the "magnitude" (for example the absolute value in the univariate case) of $u = Q^{-1}(y)$ informs us about the order of the quantile: if u is close to -1 (resp. to +1), y is a quantile with order p close to 0 (resp. to 1).
3) We have introduced the characterization $Q(u)$ for the quantile because it can be generalized in the multivariate framework. In practice, we will use this characterization to calculate the estimator of the quantile.

2.3 Estimation

Let $Y_1, \ldots, Y_n$ be n observations of y in $\mathbb{R}$. A nonparametric estimator of the *c.d.f* F is given, for $y \in \mathbb{R}$, by:

$$F_n(y) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{Y_i \le y\}}.$$

Thus, for $p \in (0,1)$, we can deduce an estimator $Q_{F_n}(p)$ of $Q_F(p)$ as follows:

$$Q_{F_n}(p) = F_n^{-1}(p) = \inf\{y : F_n(y) \ge p\}.$$

For $u = 2p-1$, using charaterization given in (2), the estimator $Q_n(u)$ of $Q(u)$ can be viewed as the solution y of the following equation

$$\frac{1}{n}\sum_{i=1}^{n} S(y - Y_i) = u. \quad (4)$$

It is easy to show that $Q_n(u) = Q_{F_n}(\frac{1+u}{2}) = Q_{F_n}(p)$ is an estimator of the quantile $Q(u) = Q_F(p)$. In fact, we have

$$F_n\left(F_n^{-1}(p)\right) - p = \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{1}_{\{Y_i \le F_n^{-1}(p)\}} - p\right) = \frac{1}{2n}\sum_{i=1}^{n}\left[S\left(Q_n(u) - Y_i\right) - u\right],$$

Using the charaterization (3) given by the minimization approach and for $u = 2p - 1$, the quantile $Q_M(p)$ can be estimated by

$$Q_{M,n}(u) = \arg\min_{\theta\in\mathbb{R}} \sum_{i=1}^{n} \phi\left(u, Y_i - \theta\right) = \arg\min_{\theta\in\mathbb{R}} \sum_{i=1}^{n} |Y_i - \theta| + u\left(Y_i - \theta\right).$$

It is easy to check that, for $u = 2p - 1$, the estimator $Q_{M,n}(u)$ of the quantile can be represented as the solution y of the equation (4). Thus, for $u = 2p - 1$, these estimators of the quantile are equal: $Q_{F_n}(p) = Q_n(u) = Q_{M,n}(u)$

3 Spatial quantile

When the random variable $\mathbf{Y}$ is a vector of $\mathbb{R}^d$, the definition of univariate quantile given by equation (1) is not valid because it is based on the idea to order the observations. However, in $\mathbb{R}^d$, the order is not total.

From now on, the vectors are considered as column and the superscript "T" is used to indicate the transpose of vectors or matrices. We suppose that $\mathbf{Y} \in \mathbb{R}^d$. In the statistical literature, multivariate quantiles have been studied by a certain number of authors, see for example Abdous and Theodorescu (1992) and Chaudhuri (1996). We choose here to focus on the approach proposed by Chaudhuri.

3.1 Two characterizations of spatial quantile

3.2 Characterization by equation root

Let $\mathbf{S}$ be a function defined as $\mathbf{S}\left(\mathbf{v}\right) = \dfrac{\mathbf{v}}{||\mathbf{v}||}$ for any non null vector $\mathbf{v} \in \mathbb{R}^d$. Let $\mathbf{u}$ be a vector of the unit ball $B^d = \left\{\mathbf{u} \in \mathbb{R}^d : ||\mathbf{u}|| < 1\right\}$. If $\mathbf{Y}$ is an absolutely continuous random variable, $\mathbf{Q}\left(\mathbf{u}\right)$ is the unique solution $\mathbf{y}$ of the following equation:

$$E\left(\mathbf{S}\left(\mathbf{y} - \mathbf{Y}\right)\right) - \mathbf{u} = 0. \tag{5}$$

For any $\mathbf{y} \in \mathbb{R}^d$, we can calculate the corresponding vector $\mathbf{u} \in B^d$ by

$$\mathbf{Q}^{-1}\left(\mathbf{y}\right) = E\left(\mathbf{S}\left(\mathbf{y} - \mathbf{Y}\right)\right).$$

3.3 Characterization by minimization approach

According to Chaudhuri (1996), the definition of the spatial quantile is a generalization of the univariate quantile definition introduced by Koenker and Basset (1978) and given by the equation (3). We consider the *multivariate* loss function defined as

$$\phi\left(\mathbf{u}, \mathbf{t}\right) = ||\mathbf{t}|| + \left\langle \mathbf{u}, \mathbf{t} \right\rangle,$$

where $||.||$ is the usual Euclidean norm and $< .,. >$ is the usual Euclidean inner product, with $\mathbf{t} \in \mathbb{R}^d$ and $\mathbf{u} \in B^d$.

Chaudhuri proposed to define the spatial quantile as follows:

$$\mathbf{Q}_M(\mathbf{u}) = \arg\min_{\theta \in \mathbb{R}^d} E\{\phi(\mathbf{u}, \mathbf{Y} - \theta)\}.$$

The function $E\{\phi(\mathbf{u}, \mathbf{Y} - \theta)\}$ is defined only when $E(\mathbf{Y}) < \infty$. Using an artifice of Kemperman (1987), the function $E\{\phi(\mathbf{u}, \mathbf{Y} - \theta) - \phi(\mathbf{u}, \mathbf{Y})\}$ is always defined. These two functions admit the same minimum when this one exists. This makes it possible to define the quantile as follows:

$$\mathbf{Q}_M(\mathbf{u}) = \arg\min_{\theta \in \mathbb{R}^d} E\{\phi(\mathbf{u}, \mathbf{Y} - \theta) - \phi(\mathbf{u}, \mathbf{Y})\}. \tag{6}$$

In a similar way to the univariate case, it is easy to check that, for any vector $\mathbf{u} \in B^d$, $\mathbf{Q}_M(\mathbf{u})$ is the solution $\mathbf{y}$ of the equation (5) and therefore $\mathbf{Q}_M(\mathbf{u}) = \mathbf{Q}(\mathbf{u})$.

3.4 Estimation

Let F_n be an empirical nonparametric estimator of F obtained from the observations $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ of $\mathbf{Y} \in \mathbb{R}^d$. We can define an estimator $\mathbf{Q}_n(.)$ of the spatial quantile $\mathbf{Q}(.)$ for all $\mathbf{u} \in B^d$, by:

$$\begin{aligned} \mathbf{Q}_n(\mathbf{u}) &= \arg\min_{\theta \in \mathbb{R}^d} \int (\phi(\mathbf{u}, \mathbf{y} - \theta) - \phi(\mathbf{u}, \mathbf{y})) F_n(d\mathbf{y}) \\ &= \arg\min_{\theta \in \mathbb{R}^d} \sum_{i=1}^{n} (\phi(\mathbf{u}, \mathbf{Y}_i - \theta) - \phi(\mathbf{u}, \mathbf{Y}_i)) \end{aligned}$$

The vector $\mathbf{u}$ gives us information about the estimator of the quantile $\mathbf{Q}_n(\mathbf{u})$. In fact,

- to determine the order of the spatial quantile, we have just to calculate the norm of $\mathbf{u}$: if $||\mathbf{u}|| \approx 1$ (resp. 0), then $\mathbf{Q}_n(\mathbf{u})$ is an extreme quantile (resp. central quantile, i.e. close to the spatial median).

- $\mathbf{u}$ is a vector of B^d, its direction indicates the position of the spatial quantile compared to the spatial median.

From the characterizations 3.2 and 3.3, it is easy to chek that, for $\mathbf{u} \in B^d$, the estimator $\mathbf{Q}_n(\mathbf{u})$ of the spatial quantile $\mathbf{Q}(\mathbf{u})$ can be seen as the solution $\mathbf{y}$ of the following equation:

$$\int \mathbf{S}(\mathbf{y} - \mathbf{t}) F_n(d\mathbf{t}) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{S}(\mathbf{y} - \mathbf{Y}_i) = \mathbf{u}. \tag{7}$$

Remarks.

- The term $||\mathbf{u}||$ said “extent of deviation” must not be considered as the Euclidean distance between $\mathbf{Q}(\mathbf{u})$ and the spatial median $\mathbf{M} = \mathbf{Q}(0)$. Moreover, the distance between $\mathbf{Q}(\mathbf{u})$ and $\mathbf{M}$ does not increase with $||\mathbf{u}||$.

- Contrary to the univariate case where $u = 2p - 1$, the "magnitude" $||\mathbf{u}||$ does not carry any probabilistic interpretation where $d \geq 2$. In particular, let us consider the region $\{\mathbf{Q}_n(\mathbf{u}) : ||\mathbf{u}|| \leq 0.5\}$. In the univariate case, it corresponds to the interquartile region with $\frac{1}{4} \leq p \leq \frac{3}{4}$. In the multivariate case, this region does not necessarily contain 50% of observations.

These two remarks are illustrated below by two examples inspired from Serfling (2002).

Example 1. Let $F = \frac{1}{2}F_1 + \frac{1}{2}F_2$, with F_1 and F_2 two uniform distributions respectively on $[-100, 0]$ and $[0, 1]$. The following quantiles are calculated: $M = 0$, $Q\left(\frac{1}{2}\right) = Q_F\left(\frac{3}{4}\right) = \frac{1}{2}$, $Q\left(-\frac{1}{2}\right) = Q_F\left(\frac{1}{4}\right) = -50$ and $Q(-0.1) = Q_F(0.45) = -10$.

- For $u = \pm\frac{1}{2}$, we have $|u| = \frac{1}{2}$ but the corresponding quantiles $Q(\frac{1}{2}) = \frac{1}{2}$ and $Q(-\frac{1}{2}) = -50$ are not equidistant compared to the median.

- For $u_1 = -0.1$ and $u_2 = \frac{1}{2}$ we have $|u_1| < |u_2|$ but $|Q(-0.1)| > |Q\left(\frac{1}{2}\right)|$. We observe here that the Euclidean distance between the quantile and the median does not increase with $|u|$.

Example 2. We consider 12 points, $\{\mathbf{y}_1, \ldots, \mathbf{y}_{12}\}$ in $\mathbb{R}^2$ given in Table 1. We give for every observation a quantile interpretation, $\mathbf{y}_i = \mathbf{Q}(\mathbf{u}_i)$, then we calculate, using the equation (7), the vector $\mathbf{u}_i = \frac{1}{n}\sum_{j=1}^{12} \mathbf{S}(\mathbf{y}_i - \mathbf{y}_j)$, and its norm $||\mathbf{u}_i||$. These two quantities are specified in Table 1.

i	$\mathbf{y}_i = \mathbf{Q}(\mathbf{u}_i)$	$\mathbf{u}_i$	$\|\|\mathbf{u}_i\|\|$
1	(0,1)	(0.011, 0.251)	0.252
2	(0,-1)	(0.011, -0.252)	0.252
3	(1,0)	(0.273, 0.000)	0.273
4	(-1,0)	(- 0.273, 0.000)	0.273
5	(0,3)	(0.039, 0.505)	0.5060
6	(0,-3.1)	(0.039, - 0.505)	0.5079
7	(0,15)	(0.368, 0.735)	0.736
8	(0,-15)	(0.368, - 0.735)	0.736
9	(0,20)	(0.030, 0.907)	0.908
10	(0,-20)	(0.030, - 0.907)	0.908
11	(-10,0)	(0.825, 0.000)	0.742
12	(1.7,0)	(0.507, 0.000)	0.5077

Table 1: Data points $\mathbf{y}_i$, values of the corresponding vectors $\mathbf{u}_i$ and their norms $||\mathbf{u}_i||$ used in Example 2. (The various values of $\mathbf{u}_i$ and $||\mathbf{u}_i||$ were round with the thousandths.)

The observations which are in the region $\{\mathbf{Q}(\mathbf{u}) : ||\mathbf{u}|| \leq 0.5\}$ are here the four points $\mathbf{y}_1, \ldots, \mathbf{y}_4$ which represent only the one third of the observations and not the half one's.

In the following paragraph, we recall the algorithm of Chaudhuri (1996) allowing to obtain an estimator of the spatial quantile.

3.5 Algorithm

The computation of the spatial median as being the quantity $\mathbf{M}$ that minimize $\sum_{i=1}^{n} ||\mathbf{Y}_i - \mathbf{M}||$ was approached by Bedall and Zimmermann (1979) and Gower (1974). Minimization algorithms were proposed by these authors. Recently, Chaudhuri (1996) proposed an iterative algorithm allowing to calculate the estimator of the spatial quantile corresponding to a fixed direction $\mathbf{u}$. This algorithm is based on the following result.

Theorem 3.1 *Let $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n$ with $\boldsymbol{Y}_i \in \mathbb{R}^d$ be a sample of distinct observations of $\mathbb{R}^d$. Let $\boldsymbol{Q}_n(\boldsymbol{u})$ be an estimator of the spatial quantile $\boldsymbol{Q}(\boldsymbol{u})$.*

- If $\boldsymbol{Q}_n(\boldsymbol{u}) \neq \boldsymbol{Y}_i, \quad \forall\ 1 \leq i \leq n$, then

$$\sum_{i=1}^{n} \frac{\boldsymbol{Y}_i - \boldsymbol{Q}_n(\boldsymbol{u})}{||\,\boldsymbol{Y}_i - \boldsymbol{Q}_n(\boldsymbol{u})||} + n\boldsymbol{u} = 0.$$

- If $\exists\ 1 \leq i \leq n$ such as $\boldsymbol{Q}_n(\boldsymbol{u}) = \boldsymbol{Y}_i$, then

$$\left\| \sum_{1 \leq i \leq n;\, \boldsymbol{Y}_i \neq \boldsymbol{Q}_n(\boldsymbol{u})} \left[\frac{\boldsymbol{Y}_i - \boldsymbol{Q}_n(\boldsymbol{u})}{||\,\boldsymbol{Y}_i - \boldsymbol{Q}_n(\boldsymbol{u})||} + \boldsymbol{u} \right] \right\| \leq \sum_{1 \leq i \leq n;\, \boldsymbol{Y}_i = \boldsymbol{Q}_n(\boldsymbol{u})} (1 + ||\boldsymbol{u}||).$$

The proof of this theorem is detailed in the article of Chaudhuri (1996). Then the corresponding algorithm of Chaudhuri (1996) comprises two steps:

- ***Step 1.*** For each $1 \leq i \leq n$, we test the following condition:

$$\left\| \sum_{1 \leq j \leq n; j \neq i} \left[\frac{\mathbf{Y}_j - \mathbf{Y}_i}{||\mathbf{Y}_j - \mathbf{Y}_i||} \right] + (n-1)\mathbf{u} \right\| \leq (1 + ||\mathbf{u}||). \quad (8)$$

  If this condition is satisfied for some i, then $\mathbf{Q}_n(\mathbf{u}) = Y_i$.
  Otherwise, one moves to the next step and tries to solve the following equation:

$$\sum_{i=1}^{n} \frac{\mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u})}{||\mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u})||} + n\mathbf{u} = 0. \quad (9)$$

- ***Step 2.*** This step consist to resolve, with an iterative way, the equation (9). Let us denote by $\mathbf{Q}_n^{(1)}(\mathbf{u})$ an initial approximation of $\mathbf{Q}_n(\mathbf{u})$. In practice we can choose, for $\mathbf{Q}_n^{(1)}(\mathbf{u})$, the vector of empirical marginal medians of the d components of $\mathbf{Y}$, calculated from the observations $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$.
  Let $\mathbf{Q}_n^{(1)}(\mathbf{u}), \ldots, \mathbf{Q}_n^{(m)}(\mathbf{u})$ be successive approximations of $\mathbf{Q}_n(\mathbf{u})$ obtained from the first m iterations. The $(m+1)^{th}$ approximation is computed in the following way.
  Let

$$\Delta = \sum_{i=1}^{n} \frac{\mathbf{Y}_i - \mathbf{Q}_n^{(m)}(\mathbf{u})}{||\mathbf{Y}_i - \mathbf{Q}_n^{m}(\mathbf{u})||} + n\mathbf{u},$$

  and

$$\Phi = \sum_{i=1}^{n} \frac{1}{||\mathbf{Y}_i - \mathbf{Q}_n^{(m)}(\mathbf{u})||} \left[I_d - \frac{(\mathbf{Y}_i - \mathbf{Q}_n^{(m)}(\mathbf{u}))(\mathbf{Y}_i - \mathbf{Q}_n^{(m)}(\mathbf{u}))^T}{||\mathbf{Y}_i - \mathbf{Q}_n^{(m)}(\mathbf{u})||^2} \right],$$

where I_d is the $d \times d$ identity matrix. When the observations $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ are not lied on a single straight line, the matrix Φ is positive definite, and in this case, one defines:

$$\mathbf{Q}_n^{(m+1)}(\mathbf{u}) = \mathbf{Q}_n^{(m)}(\mathbf{u}) + \Phi^{-1}\Delta.$$

In practice, we stop iterations when one obtains two closely successive approximations.

4 Conditional spatial quantile

We generalize in this section the previous results in the conditional framework.

4.1 Definition

Having a sample of observations $\{(\mathbf{X}_1, \mathbf{Y}_1), \ldots, (\mathbf{X}_n, \mathbf{Y}_n)\}$ from a vector $(\mathbf{X},\mathbf{Y})$ with values in $\mathbb{R}^s \times \mathbb{R}^d$, we are interested in studying the relationship between $\mathbf{X}$ and $\mathbf{Y}$. The conditional quantiles represent a mean to approach this problem.
In the univariate case (i.e. $\mathbf{Y} \in \mathbb{R}$), when the functionnal form between $\mathbf{X}$ and $\mathbf{Y}$ is unknown, there is a large variety of methods allowing to estimate conditional quantiles. For example we quote the kernel estimation, the local constant kernel estimation and the double kernel estimation (see Gannoun et al. (2002) for a description of these methods). On the other hand, few authors are interested in the conditional spatial quantile and their properties. Recently De Gooijer et al. (2006) have introduced the conditional spatial quantile based on the minimization of the pseudo-norm given by Abdous and Theodorescu (1992).
We present here an alternative formalization of the conditional spatial quantile based on generalization of the notion of spatial quantile studied by Chaudhuri (1996). Chaudhuri indexes the spatial quantile by a vector $\mathbf{u}$ in B^d which allows to give us not only the idea about the "extreme" and "central" observations, but also about their position in the multivariate scatterplots.
We define the conditional spatial quantile of the variable $\mathbf{Y}$ given $\mathbf{X} = \mathbf{x}$ as:

$$\mathbf{Q}(\mathbf{u}|\mathbf{x}) = \arg\min_{\theta \in \mathbb{R}^d} \int_{\mathbb{R}^d} \{\phi(\mathbf{u}, \mathbf{y} - \theta) - \phi(\mathbf{u}, \mathbf{y})\} F(d\mathbf{y}|\mathbf{x}). \tag{10}$$

Moreover, as in the previous section, the conditional spatial quantile can be seen as the solution $\mathbf{y}$ of the following equation:

$$E\left(\mathbf{S}\left(\mathbf{y} - \mathbf{Y}\right) \mid \mathbf{X} = \mathbf{x}\right) = \mathbf{u}. \tag{11}$$

4.2 Estimation

Let $F_n(.|\mathbf{x})$ be the nonparametric (Nadaraya-Watson) estimator of the conditional distribution function of $\mathbf{Y}$ given $\mathbf{X} = \mathbf{x}$, defined, for all $\mathbf{y} \in \mathbb{R}^d$, as

$$F_n(\mathbf{y}|\mathbf{x}) = \sum_{i=1}^{n} w_{n,i} \mathbb{1}_{\{\mathbf{Y}_i \leq \mathbf{y}\}},$$

where $w_{n,i} = \dfrac{k\left((\mathbf{x} - \mathbf{X}_i)/h_n\right)}{\sum_{i=1}^n k\left((\mathbf{x} - \mathbf{X}_i)/h_n\right)}$ is a weight associated to $\mathbf{Y}_i$, the kernel function k is a density function and h_n (the window) is a real positive sequence such that $h_n \to 0$ as $n \to \infty$.

We can deduce using equation (10), an estimator $\mathbf{Q}_n(\mathbf{u}|\mathbf{x})$ of the conditional spatial quantile $\mathbf{Q}(\mathbf{u}|\mathbf{x})$ as:

$$\begin{aligned}\mathbf{Q}_n(\mathbf{u}|\mathbf{x}) &= \arg\min_{\theta\in\mathbb{R}^d} \int_{\mathbb{R}^d} \{\phi(\mathbf{u}, \mathbf{y} - \theta) - \phi(\mathbf{u}, \mathbf{y})\} F_n(d\mathbf{y}|\mathbf{x}) \\ &= \arg\min_{\theta\in\mathbb{R}^d} \sum_{i=1}^n w_{n,i} \{\phi(\mathbf{u}, \mathbf{Y}_i - \theta) - \phi(\mathbf{u}, \mathbf{Y}_i)\}.\end{aligned}$$

From the equation (11), the estimator $\mathbf{Q}_n(\mathbf{u}|\mathbf{x})$ of the quantile $\mathbf{Q}(\mathbf{u}|\mathbf{x})$ can be viewed as the solution $\mathbf{y}$ of the following equation,

$$\int \mathbf{S}\left(\mathbf{y} - \mathbf{t}\right) F_n(d\mathbf{t}|\mathbf{x}) = \sum_{i=1}^n \mathbf{S}\left(\mathbf{y} - \mathbf{Y}_i\right) w_{n,i} = \mathbf{u}. \tag{12}$$

In the following paragraph we propose an algorithm allowing to compute an estimator of the conditional spatial quantile.

4.3 An algorithm to estimate the conditional spatial quantile

We first generalize Theorem 3.1 in the conditional case.

Theorem 4.1 *We consider n observations of couples of random vectors $\{(\boldsymbol{X}_1, \boldsymbol{Y}_1), \ldots, (\boldsymbol{X}_n, \boldsymbol{Y}_n)\}$ with values in $\mathbb{R}^s \times \mathbb{R}^d$. Let $n \geq d + s$. Let $\boldsymbol{Q}_n(\boldsymbol{u}|\boldsymbol{x})$ be an estimator of $\boldsymbol{Q}(\boldsymbol{u}|\boldsymbol{x})$.*

- *If for each $1 \leq i \leq n$, $\boldsymbol{Q}_n(\boldsymbol{u}|\boldsymbol{x}) \neq \boldsymbol{Y}_i$, then we have:*

$$\sum_{i=1}^n \frac{\boldsymbol{Y}_i - \boldsymbol{Q}_n(\boldsymbol{u}|\boldsymbol{x})}{||\,\boldsymbol{Y}_i - \boldsymbol{Q}_n(\boldsymbol{u}|\boldsymbol{x})||} K\left(\frac{\boldsymbol{x} - \boldsymbol{X}_i}{h_n}\right) + \boldsymbol{u} \sum_{i=1}^n K\left(\frac{\boldsymbol{x} - \boldsymbol{X}_i}{h_n}\right) = 0. \tag{13}$$

- *If for some i, we have $\boldsymbol{Q}_n(\boldsymbol{u}|\boldsymbol{x}) = \boldsymbol{Y}_i$, then*

$$\begin{aligned}&\left|\left| \sum_{1\leq i\leq n;\, \boldsymbol{Q}_n(\boldsymbol{u}|\boldsymbol{x})\neq \boldsymbol{Y}_i} \frac{\boldsymbol{Y}_i - \boldsymbol{Q}_n(\boldsymbol{u}|\boldsymbol{x})}{||\,\boldsymbol{Y}_i - \boldsymbol{Q}_n(\boldsymbol{u}|\boldsymbol{x})||} K\left(\frac{\boldsymbol{x} - \boldsymbol{X}_i}{h_n}\right) + \sum_{1\leq i\leq n;\, \boldsymbol{Q}_n(\boldsymbol{u}|\boldsymbol{x})\neq \boldsymbol{Y}_i} K\left(\frac{\boldsymbol{x} - \boldsymbol{X}_i}{h_n}\right) \boldsymbol{u} \right|\right| \\ &\qquad \leq \sum_{1\leq i\leq n;\, \boldsymbol{Q}_n(\boldsymbol{u}|\boldsymbol{x})= \boldsymbol{Y}_i} K\left(\frac{\boldsymbol{x} - \boldsymbol{X}_i}{h_n}\right)(1 + ||\boldsymbol{u}||).\end{aligned} \tag{14}$$

The proof of this theorem is postponed to the Appendix. Using this theorem, the algorithm to compute the estimator of the conditional spatial quantile splits into two steps.

- ***Step 1.*** For each $1 \leq i \leq n$, we test the following inequality:

$$\left|\left| \sum_{1\leq j\leq n;\ j\neq i} \frac{\mathbf{Y}_j - \mathbf{Y}_i}{||\mathbf{Y}_j - \mathbf{Y}_i||} K\left(\frac{\mathbf{x} - \mathbf{X}_j}{h_n}\right) + \sum_{1\leq j\leq n;\ i\neq j} K\left(\frac{\mathbf{x} - \mathbf{X}_j}{h_n}\right) \mathbf{u} \right|\right|$$

$$\leq K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right)(1+||\mathbf{u}||). \tag{15}$$

If this condition is satisfied for the observation i, then $\mathbf{Q}_n(\mathbf{u}|\mathbf{x}) = \mathbf{Y}_i$.

Otherwise one passes to the second step which consists in resolving numerically equation (13).

- ***Step 2.*** Let the initial approximation $\mathbf{Q}_n^{(1)}(\mathbf{u}|\mathbf{x})$ ($\in \mathbb{R}^d$) be the vector of the empirical conditional medians of $\mathbf{Y}$, computed from the observations $\{(\mathbf{X}_1, \mathbf{Y}_1), \ldots, (\mathbf{X}_n, \mathbf{Y}_n)\}$. We denote by $\mathbf{Q}_n^{(1)}(\mathbf{u}|\mathbf{x}), ..., \mathbf{Q}_n^{(m)}(\mathbf{u}|\mathbf{x})$ successive approximations of $\mathbf{Q}_n(\mathbf{u}|\mathbf{x})$

  The $(m+1)^{th}$ approximation $\mathbf{Q}_n^{(m+1)}(\mathbf{u}|\mathbf{x})$ is computed as follows.

  Let
$$\Delta = \sum_{i=1}^{n} \frac{\mathbf{Y}_i - \mathbf{Q}_n^{(m)}(\mathbf{u}|\mathbf{x})}{||\mathbf{Y}_i - \mathbf{Q}_n^{(m)}(\mathbf{u}|\mathbf{x})||}\ K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right) + \mathbf{u}\sum_{i=1}^{n} K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right)$$

  and

$$\Phi = \sum_{i=1}^{n} \frac{1}{||\mathbf{Y}_i - \mathbf{Q}_n^{(m)}(\mathbf{u}|\mathbf{x})||}\left[I_d - \frac{(\mathbf{Y}_i - \mathbf{Q}_n^{(m)}(\mathbf{u}|\mathbf{x}))\ (\mathbf{Y}_i - \mathbf{Q}_n^{(m)}(\mathbf{u}|\mathbf{x}))^T}{||\mathbf{Y}_i - \mathbf{Q}_n^{(m)}(\mathbf{u}|\mathbf{x})||^2}\right]\ K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right).$$

  If the observations $\mathbf{Y}_i$ are not lied on the single straight line, then Φ is a defined positive matrix and we define:

$$\mathbf{Q}_n^{(m+1)}(\mathbf{u}|\mathbf{x}) = \mathbf{Q}_n^{(m)}(\mathbf{u}|\mathbf{x}) + \Phi^{-1}\ \Delta.$$

  Iteration is continued until two successive approximations of $\mathbf{Q}_n(\mathbf{u}|\mathbf{x})$ happen to be sufficiently close.

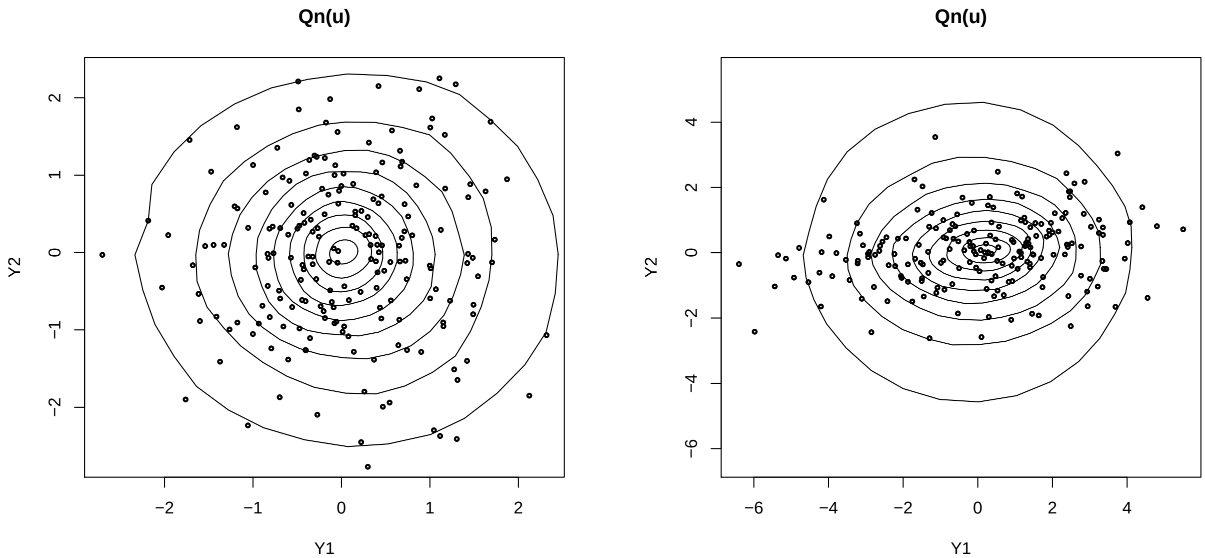
5 Simulations

In order to make easy the realization and the interpretation of the graphics, we suppose that $d = 2$ (two-dimensional case). The identification of the extreme observations in a sample represents an important step in a statistical study. In the univariate case, we can determine these values using the boxplot. In this section, we give a graphic (called quantile contour plot) which can be seen as the boxplot in the multivariate framework.
In this simulation study, we consider a vector $\mathbf{u} \in B^2$ of the form $(rcos\theta, rsin\theta)^T$ with r taking its values in $\{r_k = \frac{k}{10}, k = 1, \ldots, 9\}$ and θ taking its values in $\{\theta_l = \frac{\pi l}{16}, l = 0, 1, \ldots, 31\}$. Then we compute for each vector $\mathbf{u}$ the corresponding spatial quantile. The set $\{\mathbf{Q}_n(\mathbf{u}) : ||\mathbf{u}|| = r\}$, with $0 < r < 1$, is named "quantile contour plot". This set can be considered as the equivalent of the boxplot in the multivariate case (see Chakraborty (2001)). When the norm r of $\mathbf{u}$ is close to 1, the observations located outside this contour can be classed as extreme. The choice of r depends on the study framework. Generally, the specialist fixes it according to its objectives.

5.1 A first simulation: case of unconditional spatial quantiles

To illustrate the construction of quantile contour plot, we simulate 200 observations according to the multinormal distribution $N_2(0, I_2)$. We note by Y_1 and Y_2 the two components of $\mathbf{Y} \in \mathbb{R}^2$.

In order to compute the quantile contour plot of radius r, we use the vector $\mathbf{u}$ such that $||\mathbf{u}|| = r$ while the angle θ varies from θ_0 to θ_{31}. Then we interpolate the estimated spatial quantiles in order to get the corresponding quantile contour plot. Figure 1 (a) represent nine estimated contours (from 10% to 90%) ploted on the corresponding scatterplot.



(a) (b)

Figure 1: Quantile contour plots from 10% to 90% for observations given by (a) a $N_2(0, I_2)$ distribution, and (b) by a $N_2(0, \Sigma)$ distribution.

In order to make sure that contours adapt with the form of the scatter plot, we simulate 200 observations according to the multivariate normal distribution $N_2(0, \Sigma)$ with $\Sigma = \begin{pmatrix} 5 & 0.4 \\ 0.4 & 1 \end{pmatrix}$. Figure 1 (b) shows that the contours have a different form than those presented in Figure 1 (a), this confirms that they take well into account the various variances and covariances.

5.2 A second simulation: case of conditional spatial quantiles

In order to see the behavior of the conditional spatial quantile estimators, while varying the vector $\mathbf{u}$, we have simulated 200 observations according to the following multivariate normal distribution:

$$\begin{pmatrix} Y_1 \\ Y_2 \\ X \end{pmatrix} \sim N_3 \left(0, \Gamma = \begin{pmatrix} 5 & 0.2 & 0.4 \\ 0.2 & 1 & 0.9 \\ 0.4 & 0.9 & 1 \end{pmatrix} \right).$$

In this example, we have fixed $x = 0$. Then for each value of $\mathbf{u}$, we compute the estimator of the corresponding conditional spatial quantile.

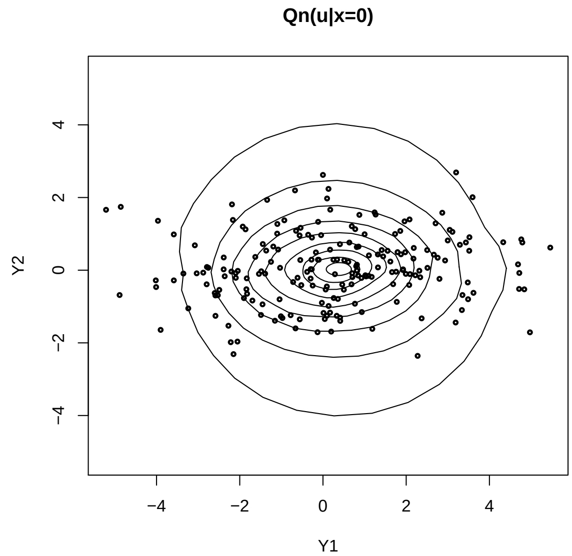


Figure 2: Conditional quantile contour plots from 10% to 90% for $x = 0$

Figure 2 shows that the conditional quantile contour plots from 10% to 90% ploted using the estimators of the conditional spatial quantiles adapt well with the form of the scatterplot. In addition, we know that the estimator of the spatial median (corresponding to $\mathbf{u} = (0, 0)$) converges asymtotically to the true median which is here for a multivariate normal distribution equal to the mean $(0, 0)$, so to check the quality of the estimator we have compared the estimated spatial median to the theoretical mean. For $\mathbf{u} = (0, 0)$, we have $\mathbf{Q}_n(\mathbf{u}\ |x = 0) = (0.08, -0.03)$, which is very close to $(0, 0)$.

Appendix

Proof of Theorem 4.1

- The first result can be deduced directly from the equation (11). If the observations are not lying in a single straight line in $\mathbb{R}^d$, then the conditional spatial quantile is the unique solution $\mathbf{y}$ of the equation (11). Then we deduce that $\mathbf{Q}_n(\mathbf{u}|\mathbf{x})$ satisfy the following equation:

$$\sum_{i=1}^{n} \frac{\mathbf{Q}_n(\mathbf{u}|\mathbf{x}) - \mathbf{y}}{||\mathbf{Q}_n(\mathbf{u}|\mathbf{x}) - \mathbf{y}||} K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right) = \mathbf{u} \sum_{i=1}^{n} K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right).$$

- Let us prove now the second part of the theorem. The function $\Phi(\mathbf{u}, \mathbf{y})$ is a convex function on $\mathbb{R}^d$ and it depends on $\mathbf{y}$. One deduces that

$$\mathbf{Q}_n(\mathbf{u}|\mathbf{x}) = \arg\min_{\mathbf{Q}} \sum_{i=1}^{n} \Phi(\mathbf{u}, \mathbf{Y}_i - \mathbf{Q}) K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right)$$

if and only if, for any $\mathbf{h} \in \mathbb{R}^d$, we have

$$\lim_{t \to 0^+} \left[\sum_{i=1}^{n} \Phi(\mathbf{u}, \mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u}|\mathbf{x}) + t\mathbf{h}) K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right) - \sum_{i=1}^{n} \Phi(\mathbf{u}, \mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u}|\mathbf{x})) K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right)\right] \geq 0.$$

However, for all $\mathbf{y}$, $\mathbf{h} \in \mathbb{R}^d$ such as $\mathbf{y} \neq 0$, we get:

$$\lim_{t \to 0^+} \frac{\Phi(\mathbf{u}, \mathbf{y} + t\mathbf{h}) - \Phi(\mathbf{u}, \mathbf{y})}{t} = \lim_{t \to 0^+} \frac{||\mathbf{y} + t\mathbf{h}|| - ||\mathbf{y}|| + < \mathbf{u}, t\mathbf{h} >}{t} \quad = \quad < \frac{\mathbf{y}}{||\mathbf{y}||} + \mathbf{u}, \mathbf{h} > .$$

Moreover, for all $\mathbf{h} \in \mathbb{R}^d$ and $\mathbf{y} = 0$, we have

$$\lim_{t \to 0^+} \frac{\Phi(\mathbf{u}, t\mathbf{h}) - \Phi(\mathbf{u}, 0)}{t} = || \ \mathbf{h} \ || + < \mathbf{u}, \mathbf{h} > .$$

Thereafter, using those two properties on the previous inequality, we obtain:

$$\sum_{1 \leq i \leq n; \mathbf{Q}_n(\mathbf{u}|\mathbf{x}) \neq \mathbf{Y}_i} K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right) < \frac{\mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u}|\mathbf{x})}{||\mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u}|\mathbf{x})||} + \mathbf{u}, \mathbf{h} >$$

$$+ \sum_{1 \leq i \leq n; \mathbf{Q}_n(\mathbf{u}|\mathbf{x}) = \mathbf{Y}_i} K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right) (|| \ \mathbf{h} \ || + < \mathbf{u}, \mathbf{h} >) \geq 0.$$

Because this inequality is true for all $\mathbf{h} \in \mathbb{R}^d$, it is true also for $-\mathbf{h}$. While replacing $\mathbf{h}$ par $-\mathbf{h}$ in the previous inequality, we obtain then:

$$\sum_{1 \leq i \leq n; \mathbf{Q}_n(\mathbf{u}|\mathbf{x}) = \mathbf{Y}_i} K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right) (|| \ \mathbf{h} \ || - < \mathbf{u}, \mathbf{h} >) \quad \geq$$

$$\sum_{1 \leq i \leq n; \mathbf{Q}_n(\mathbf{u}|\mathbf{x}) \neq \mathbf{Y}_i} K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right) < \frac{\mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u}|\mathbf{x})}{||\mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u}|\mathbf{x})||} + \mathbf{u}, \mathbf{h} > . \tag{16}$$

On the other hand, using the Schwartz inequality, we get:

$$| \ || \ \mathbf{h} \ || \pm < \mathbf{u}, \mathbf{h} >| \quad \leq || \ \mathbf{h} \ || + | < \mathbf{u}, \mathbf{h} > | \leq (1 + || \ \mathbf{u} \ || \ )|| \ \mathbf{h} \ ||.$$

Thus, inequality (16) is equivalent to

$$\sum_{1 \leq i \leq n; \mathbf{Q}_n(\mathbf{u}|\mathbf{x}) = \mathbf{Y}_i} K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right) (1 + || \ \mathbf{u} \ ||)|| \ \mathbf{h} \ || \quad \geq$$

$$\sum_{1 \leq i \leq n; \mathbf{Q}_n(\mathbf{u}|\mathbf{x}) \neq \mathbf{Y}_i} K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right) < \frac{\mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u}|\mathbf{x})}{||\mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u}|\mathbf{x})||} + \mathbf{u}, \mathbf{h} > . \tag{17}$$

Because this inequality is true for all $\mathbf{h} \in \mathbb{R}^d$, we can choose in particular

$$\mathbf{h} = \frac{\mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u}|\mathbf{x})}{||\mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u}|\mathbf{x})||} + \mathbf{u}. \tag{18}$$

and if we put this value of h in equation (17), we get :

$$\sum_{1 \leq i \leq n; \mathbf{Q}_n(\mathbf{u}|\mathbf{x}) = \mathbf{Y}_i} K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right)(1 + || \mathbf{u} ||) \quad \geq$$

$$\left\| \sum_{1 \leq i \leq n; \mathbf{Q}_n(\mathbf{u}|\mathbf{x}) \neq \mathbf{Y}_i} K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right)\left(\frac{\mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u}|\mathbf{x})}{||\mathbf{Y}_i - \mathbf{Q}_n(\mathbf{u}|\mathbf{x})||} + \mathbf{u}\right) \right\|$$

Then we deduce the inequality (14).

□

Remark. The **R**-codes allowing to estimate spatial quantiles, conditional spatial quantiles and quantiles contour plots are available and can be asked to the authors.

References

Abdous, B. and Theodorescu, R. (1992). Note on the spatial quantile of a random vector. *Statistics and Probability Letter*, **13**, 333-336.

Barnett, V. (1976). The ordering of multivariate data (with comments). *Journal of Royal Statistical Society, Ser. A*, **139**, 318-354.

Bedall, F.K. and Zimmermann, H. (1979). Algorithm AS 143, the Mediancenter. *Applied Statistics*, **28**, 325-328.

Brown, B. M., and Hettmansperger, T. P. (1987). Affine Invariant Rank Methods in the Bivariate Location Model. *Journal of the Royal Statistical Society, Ser. B*, **49**, 301-310.

Brown, B. M., and Hettmansperger, T. P. (1989). An Affine Invariant Bivariate Versions of the Sign Test. *Journal of the Royal Statistical Society, Ser. B*, **51**, 117-125.

Chakraborty, B. (2001). On affine equivariant multivariate quantiles. *The Institute of Statistical Mathematics*, **53**, 380-403.

Chaudhuri, P. (1992a). Multivariate location estimation using extension of R-estimates through U-statistics type approach. *Annals of Statistics*, **20**, 897-916.

Chaudhuri, P. (1992b). Generalized regression quantiles: Forming a useful toolkit for robust linear regression. L_1 *Statistical Analysis and Related Methods*, ed. Y. Dodge, Amesterdam: North-Holland, pp. 169-185.

Chaudhuri, P. (1996). On a geometric notation of quantiles for multivariate data. *Journal of the American Statistical Association*, **91**, 862-872.

De Gooijer, J. G., Gannoun, A. and Zerom, D. (2006). A multivariate quantile predictor. *Communications in Statistics-Theory and Methods*, **35**, 133-147.

Donoho, D. L. and Gasko, M. (1992). Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *Annals of Statistics*, **20**, 1803-1827.

Eddy, W.F. (1985). Ordering of Multivariate Data. *Computer Science and Statistics: The Interface*, ed. L. Billard, Amesterdam: North-Holland, 25-30.

Efron, B. (1991). Regression percentiles using asymmetric squared error loss. *Statistica Sinica*, **1**, 93-125.

Ferguson, T. (1967). *Mathematical Statistics: A Decision Theoric Approach*. Academic Press: New York.

Gannoun, A., Girard, S., Guinot, C. and Saracco, J. (2002). Reference curves based on nonparametric quantile regrassion. *Statistics in Medecine*, **21**, 3119-3135.

Gower, J. C. (1974). Algorithm AS 78: The Mediancenter. *Applied Statistics*, **23**, 466-470.

Koenker, R. and Basset, G. (1978). Regression Qantiles. *Econometrica*, **46**, 33-50.

Koenker, R. and Portnoy, S. (1987). L Estimation for linear models. *Journal of the American statistical Association*, **82**, 851-857.

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

Liu, R. Y., Parelius, J. M. and Singh, K. (1999). Multivariate analysis by data depth: Descriptive statistics, graphics and inference (with discussion). *Annals of Statistics*, **27**, 783-858.

Oja, H. (1983). Descriptive Statistics for Multivariate Trimming. *Statistics and Probability Letters*, **1**, 327-332.

Plackett, R. L. (1976). Comment on Ordering of multivariate data by V. Barnett. *Journal of the Royal statistical Society, Ser.A*, **139**, 344-346.

Reiss, R. D. (1989). Approximation distributions of order statistics with applications to nonparametric statistics. New York: Springer.

Serfling, R. (1980). Approximation theorem of mathematical statistics. New York: John Wiley.

Serfling, R. (2002). Quantile functions for multivariate analysis: approaches and applications. *Annals of Statistics*, **25**, 435-477.

Cahiers du GREThA
Working papers of GREThA

GREThA UMR CNRS 5113

Université Montesquieu Bordeaux IV
Avenue Léon Duguit
33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75
Fax : +33 (0)5.56.84.86.47

www.gretha.fr

Cahiers du GREThA (derniers numéros)

2007-17 : BERTIN Alexandre, LEYLE David, *Mesurer la pauvreté multidimensionnelle dans un pays en développement Démarche méthodologique et mesures appliquées au cas de l'Observatoire de Guinée Maritime*
2007-18 : DOUAI Ali, *Wealth, Well-being and Value(s): A Proposition of Structuring Concepts for a (real) Transdisciplinary Dialogue within Ecological Economics*
2007-19 : AYADI Mohamed, RAHMOUNI Mohieddine, YILDIZOGLU Murat, *Sectoral patterns of innovation in a developing country: The Tunisian case*
2007-20 : BONIN Hubert, *French investment banking at Belle Epoque: the legacy of the 19th century Haute Banque*
2007-21 : GONDARD-DELCROIX Claire, *Une étude régionalisée des dynamiques de pauvreté Régularités et spécificités au sein du milieu rural malgache*
2007-22 : BONIN Hubert, *Jacques Laffitte banquier d'affaires sans créer de modèle de banque d'affaires (des années 1810 aux années 1840)*
2008-01 : BERR Eric, *Keynes and the Post Keynesians on Sustainable Development*
2008-02 : NICET-CHENAF Dalila, *Les accords de Barcelone permettent- ils une convergence de l'économie marocaine ?*
2008-03 : CORIS Marie, *The Coordination Issues of Relocations? How Proximity Still Matters in Location of Software Development Activities*
2008-04 : BERR Eric, *Quel développement pour le 21ème siècle ? Réflexions autour du concept de soutenabilité du développement*
2008-05 : DUPUY Claude, LAVIGNE Stéphanie, *Investment behaviors of the key actors in capitalism : when geography matters*
2008-06 : MOYES Patrick, *La mesure de la pauvreté en économie*
2008-07 : POUYANNE Guillaume, *Théorie économique de l'urbanisation discontinue*
2008-08 : LACOUR Claude, PUISSANT Sylvette, *Medium-Sized Cities and the Dynamics of Creative Services*
2008-09 : BERTIN Alexandre, *L'approche par les capabilités d'Amartya Sen, Une voie nouvelle pour le socialisme libéral*
2008-10 : CHAOUCH Mohamed, GANNOUN Ali, SARACCO Jérôme, *Conditional Spatial Quantile: Characterization and Nonparametric Estimation*

La coordination scientifique des Cahiers du GREThA est assurée par Sylvie FERRARI et Vincent FRIGANT. La mise en page est assurée par Dominique REBOLLO.

Zuo, Y. and Serfling, R. (2000). General notions of statistical depth function. *Annals of Statistics*, **28**, 461-482.