Journal of Multivariate Analysis 131 (2014) 265–278

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Maximal correlation in a non-diagonal case

F. López Blázquez *, B. Salamanca Miño

Universidad de Sevilla, Spain

ARTICLE INFO

Article history:
Received 4 March 2014
Available online 31 July 2014

AMS 2000 subject classifications:
62H20
62G30
62E10
60E15

Keywords:
Maximal correlation
Orthogonal polynomials
Tridiagonal matrices
Order statistics

ABSTRACT

We present a method for the obtention of the maximal correlation coefficient that extends the simple method given by Papadatos and Xifara (2013). We illustrate our method with the calculation of the maximal correlation between the kth largest order statistics of overlapping samples.

1. Introduction

Let (X, Y) be a random vector with non-degenerate marginal distributions. The *maximal correlation coefficient* is defined as

$$R(X, Y) := \sup \rho(g(X), h(Y)) \tag{1}$$

where $\rho(\cdot, \cdot)$ is the *Pearson correlation coefficient* and the supremum in (1) is taken over the set of measurable functions g, $h : \mathbb{R} \to \mathbb{R}$ with $0 < \operatorname{Var} g(X) < \infty$ and $0 < \operatorname{Var} h(Y) < \infty$, see Gebelein [10]. The maximal correlation is an attractive measure of dependence and it has been used in relation with transformations improving the linear dependence among the variables, see Breiman and Friedman [3] and Buja [5]. More properties and applications can be found in Lancaster [13], Liu et al. [15], Dembo et al. [9], Bryc et al. [4], and Yu [30] among others. One of the main drawbacks in the use of the maximal correlation coefficient is that in general its explicit calculation is difficult. In some cases, the maximal correlation can be obtained by using some properties of special families of orthogonal polynomials, for instance Legendre polynomials in Terrell [26], Jacobi polynomials in Székely and Móri [25], Laguerre polynomials in Nevzorov [18] and López-Blázquez and Salamanca-Miño [17] and Hahn Polynomials in López-Blázquez and Castaño Martínez [16].

From the results in Rényi [21, p. 447], it follows that if

$$R(X, Y) = |\rho(X, Y)| > 0, \tag{2}$$

then both the regressions of one variable on the other are linear, that is to say,

$$\mathbb{E}(Y \mid X) = bX + a, \quad \text{and} \quad \mathbb{E}(X \mid Y) = b'X + a', \quad \text{(a.s.)}, \tag{3}$$

* Corresponding author.
E-mail addresses: lopez@us.es (F. López Blázquez), bsm@us.es (B. Salamanca Miño).

http://dx.doi.org/10.1016/j.jmva.2014.07.008

for certain real numbers b, b', a and a'. Following Sarmanov [22,23], some authors have used the (wrong) implication (3) $\Rightarrow$ (2) claiming that it is an easy method for obtaining the maximal correlation when both regressions are linear. Unfortunately (3) $\Rightarrow$ (2) is not true as it is shown in Papadatos [19]. Recently, Papadatos and Xifara [20] provided a unified method for obtaining the maximal correlation coefficient when the bivariate distribution has a certain diagonal structure that in particular satisfies (3) along with some other assumptions. Their method is rather simple and although it requires the existence of adequate families of orthogonal polynomials, it is not necessary the explicit knowledge of them. Despite the method of these authors covers many cases of interest, it is clear that some alternative methods should be given in order to obtain the maximal correlation in non-diagonal cases.

In this work we propose a variation on the assumptions of Papadatos and Xifara [20]. Our main result is stated in Theorem 4 in which the maximal correlation is obtained as the solution of an optimization problem in infinitely many variables. A detailed study of this optimization problem is discussed in Section 2.

Our results include the diagonal case, see Theorem 4(b). Many examples of this situation are shown in Papadatos and Xifara [20]. The purpose of Section 5 is to provide an example in the non-diagonal case, so we study the maximal correlation between consecutive pairs of kth-largest order statistics obtained from the same sequence. For the sequence of partial maxima, i.e. $k = 1$, it is possible to obtain a non-trivial upper bound for the maximal correlation, see Theorem 8. This result agrees with some previous works of the authors, see Castaño Martínez et al. [7]. In the case $k > 1$, it seems that it is not possible to obtain a closed formula for the maximal correlation, nevertheless it is not difficult to obtain good numerical approximations by using the Sturm properties of some sequences of polynomials described in Section 2. The joint distribution of kth-largest order statistics in overlapping samples does not seem to be well-known, for that reason we include Section 4.

2. Preliminary results

Most of the results about tridiagonal matrices and orthogonal polynomials cited in this section can be found in Golub and Van Loan [11] and Ismail [12].

2.1. An optimization problem

Consider the spaces of real sequences $\ell^2 = \{\mathbf{w} = \{w_j\}_{j\geq 0} : \|\mathbf{w}\|^2 = \sum_{j\geq 0} w_j^2 < +\infty\}$ and for $m \geq 1$, $\ell^2_m = \{\mathbf{w} = \{w_j\}_{j\geq 0} : w_j = 0, \text{ for all } j \geq m\}$. (The space ℓ^2_m can be clearly identified to $\mathbb{R}^m$, $m \geq 1$.) Let $\{a_j\}_{j\geq 1}$ and $\{b_j\}_{j\geq 1}$ be bounded sequences of real numbers and define $T : \ell^2 \to \mathbb{R}$ as $T(\mathbf{w}) = \sum_{j\geq 1} \left(a_j w_{j-1} - b_j w_j\right)^2$. Our aim in this section is the study of the optimization problem

$$\gamma := \sup\{T(\mathbf{w}) : \mathbf{w} \in \ell^2 \text{ with } \|\mathbf{w}\| = 1\}. \tag{4}$$

Let $\gamma_m := \sup\{T(\mathbf{w}) : \mathbf{w} \in \ell^2_m \text{ with } \|\mathbf{w}\| = 1\}$. It is not difficult to check that

$$\lim_{m\to\infty} \gamma_m = \gamma.$$

Let us define the symmetric tridiagonal matrices

$$\mathbf{A}_m = \begin{pmatrix} a_1^2 & -a_1b_1 & 0 & \cdots & 0 \\ -a_1b_1 & a_2^2+b_1^2 & -a_2b_2 & \cdots & 0 \\ 0 & -a_2b_2 & a_3^2+b_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & -a_{m-1}b_{m-1} \\ 0 & 0 & 0 & \cdots & a_m^2+b_{m-1}^2 \end{pmatrix}.$$

Note that $\mathbf{A}_m = \mathbf{T}_m^t \mathbf{T}_m$ is positive semi-definite with

$$\mathbf{T}_m = \begin{pmatrix} a_1 & -b_1 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & -b_2 & 0 & \cdots & 0 \\ 0 & 0 & a_3 & -b_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & -b_{m-1} \\ 0 & 0 & 0 & 0 & \cdots & a_m \end{pmatrix}.$$

Let $\lambda_{m,m} \geq \cdots \geq \lambda_{m,1} \geq 0$ be the eigenvalues of $\mathbf{A}_m$. It is well known that $\gamma_m = \sup\{\mathbf{w}^t\mathbf{A}_m\mathbf{w} : \mathbf{w} \in \mathbb{R}^m, \mathbf{w}^t\mathbf{w} = 1\} = \lambda_{m,m}$.

If $b_j = 0$ for all $j \geq 0$, the matrix $\mathbf{A}_m$ is diagonal then $\lambda_{m,m} = \max\{a_j^2 : j = 1, \ldots, m\}$ and $\gamma = \sup\{a_m^2 : m \geq 1\}$. If there exists $j_0 \geq 1$ such that $a_{j_0}^2 = \sup\{a_m^2 : m \geq 1\}$, then the supremum in (4) is attained by $\mathbf{w}^* = \{\delta_{j,j_0}\}_{j\geq 0}$, with δ_{j,j_0} the Kronecker delta.