



Testing independence of random vectors by inverse regressions

Jean Gérard Aghoukeng Jiofack[†] and Guy Martial Nkiet[‡]

[†]Université de Yaoundé 1, Faculté des Sciences, Département de Mathématiques, BP 812 Yaoundé, Cameroun

[‡]Université des Sciences et Techniques de Masuku, Faculté des Sciences, Département de Mathématiques et Informatique, BP 943 Franceville, Gabon

Received 16 July 2012; Accepted 24 October 2012

Copyright © 2012, Journal Afrika Statistika. All rights reserved

Abstract. We propose a new test of independence of random vectors. We first show that the null hypothesis implies the nullity of the trace of an operator involving inverse regressions covariance operators. Then, using an approach based on slicing, we define a test statistic for which an asymptotic distribution under null hypothesis is derived. Simulations that permit to evaluate the performance of the proposed test with comparisons with existing methods are given.

Résumé. Nous proposons un nouveau test d'indépendance de vecteurs aléatoires. Nous exprimons tout d'abord l'hypothèse d'indépendance au moyen de la nullité de la trace d'un opérateur défini à partir des opérateurs de covariance de régressions inverses appropriées. Utilisant ensuite une approche par tranchage, nous définissons une statistique de test pour laquelle nous obtenons une loi limite sous l'hypothèse nulle d'indépendance. Cela permet de définir la méthode proposée qui est ensuite évaluée et comparée à des méthodes existantes par des simulations.

Key words: Independence test; Sliced inverse regression; Asymptotic theory.

AMS 2010 Mathematics Subject Classification : 62H15; 62J02; 62E20.

1. Introduction

Testing for independence of two random vectors $X = (X_1, \dots, X_p)'$ and $Y = (Y_1, \dots, Y_q)'$, that are respectively p -dimensional and q -dimensional, is a classical problem in statistics. When $Z = (X', Y')'$ has a $(p + q)$ -variate normal distribution with partitioned covariance

Jean Gérard Aghoukeng Jiofack : aghoukeng-maths@yahoo.fr

Guy Martial Nkiet: gnkiet@hotmail.com

matrix

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

the hypothesis of independence may be formulated as $C_{12} = 0$. In this context, several tests have been introduced, including the likelihood ratio test and the Pillai's test (see, e.g., Anderson, (1984), Bilodeau and Brenner, (1999)). For the more general case where Z has an elliptic distribution three methods have been proposed in Allaire and Lepage, (1990) and, more recently, an approach based on spatial signs have been introduced in Taskinen, *et al.* (2003). It is, of course, better to consider distribution-free methods. In this direction, Cl  roux, (1995) introduced a method based on ranks whereas nonparametric approaches have been proposed, for instance, in Bakirov, *et al.* (2006), Gieser and Randles, (1997), Meintanis and Iliopoulos, (2008), Sinha and Wieand, (1977) and Sz  kely *et al.* (2007). To the best of our knowledge, there does not exist a method that is based on a well known result in probability theory that gives connections between the independence property and conditional expectations. Such a method may be of a great interest because it is necessarily a distribution-free method since the aforementioned result holds whatever is the distribution of (X, Y) . In this paper, we tackle an approach based on this result for defining an independence test between random vectors. Our proposal is described in Section 2. We first remark that the independence property implies the nullity of the trace of an operator involving covariance operators of expectations of X conditional to the coordinates of Y . Then, we adopt ideas used in sliced inverse regression for approximating these covariance operators and, therefore, to introduce the test statistic that will be used. The limiting distribution of this statistic under null hypothesis is then derived, and the related test procedure is described. Section 3 is devoted to the presentation of simulations that permit to evaluate the performances of the proposed approach and to compare it with existing methods. All the proofs of lemmas and the theorem are given in Section 4.

2. The proposed method

This section is devoted to the presentation of our proposal for testing for independence between two random vectors. We first introduce notations and remark that the null hypothesis implies the nullity of the trace of an operator involving inverse regressions covariance operators. Then, using an approach based on slicing, as in Li, (1991), we define a test statistic for which an asymptotic distribution under null hypothesis is derived. That permits to specify the proposed testing procedure.

2.1. Formulation of the problem

Denoting by \mathbb{E} the mathematical expectation, we assume that $\mathbb{E}(\|X\|^4) < +\infty$, where $\|\cdot\|$ denotes the Euclidean norm of \mathbb{R}^p induced by the usual inner product $\langle \cdot, \cdot \rangle$ of \mathbb{R}^p . We are interested in testing for the hypothesis

$$\mathcal{H}_0 : X \perp\!\!\!\perp Y,$$

where $\perp\!\!\!\perp$ denotes stochastic independence, against the alternative hypothesis \mathcal{H}_1 stating that X and Y are not independent. If \mathcal{H}_0 is true, then $X \perp\!\!\!\perp Y_j$ for any $j \in \{1, \dots, q\}$. We

will express this latter property by means of the covariance operators Λ_j of the conditional expectations $\mathbb{E}(X|Y_j)$, given by

$$\Lambda_j = \mathbb{E}((\mathbb{E}(X|Y_j) - \mu) \otimes (\mathbb{E}(X|Y_j) - \mu)),$$

where $\mu = \mathbb{E}(X)$, and \otimes denotes the tensor product defined as follows: for any pair (x, y) of elements of an Euclidean space with inner product $\langle \cdot, \cdot \rangle$, $x \otimes y$ is the linear map $h \mapsto \langle x, h \rangle y$. Using the equality $tr(x \otimes x) = \|x\|^2$ (see Dauxois, (1994)), we obtain

$$tr(\Lambda_j) = \mathbb{E}(tr((\mathbb{E}(X|Y_j) - \mu) \otimes (\mathbb{E}(X|Y_j) - \mu))) = \mathbb{E}(\|\mathbb{E}(X|Y_j) - \mu\|^2).$$

Then $X \perp\!\!\!\perp Y_j$ implies that $\mathbb{E}(X|Y_j) = \mu$ almost surely, what is equivalent to having $tr(\Lambda_j) = 0$. Therefore, \mathcal{H}_0 implies that $tr(\Lambda) = 0$, where $\Lambda = \sum_{j=1}^q \Lambda_j$. Consequently, testing for \mathcal{H}_0 against \mathcal{H}_1 can be done by taking a consistent estimator of $tr(\Lambda)$ as test statistic. Following an approach used in Li, (1991) for estimating an inverse regression covariance operator, we will in fact use a consistent estimator of $tr(\tilde{\Lambda})$, where $\tilde{\Lambda}$ is an approximation of Λ obtained by slicing the ranges of the Y_j 's. For $j \in \{1, \dots, q\}$, let $(I_h^{(j)})_{1 \leq h \leq r_j}$ be a partition of $Y_j(\Omega)$ such that each probability $p_{jh} := P(Y_j \in I_h^{(j)})$ is non null. Putting $\mu_{jh} = \mathbb{E}(X|Y_j \in I_h^{(j)})$ and $\tau_{jh} = \mu_{jh} - \mu$, the aforementioned operator $\tilde{\Lambda}$ is given by $\tilde{\Lambda} = \sum_{j=1}^q \tilde{\Lambda}_j$, where $\tilde{\Lambda}_j = \sum_{h=1}^{r_j} p_{jh} \tau_{jh} \otimes \tau_{jh}$.

Remark 1. In all of the paper we use tensor notations and operators. However, in a finite-dimensional framework the related transcriptions into matrix notations, that are useful for practical implementation, are easy to obtain from Dauxois, (1994). More precisely, the matrix related to the operator $x \otimes y$ is given by $\mathbf{y}\mathbf{x}'$, where $\mathbf{x} = (x_1, \dots, x_p)'$ (resp. $\mathbf{y} = (y_1, \dots, y_q)'$) is the matricial representation of the vector x (resp. y) relative to the canonical basis of \mathbb{R}^p (resp. \mathbb{R}^q).

2.2. The test statistic

Letting $\{(X^{(i)}, Y^{(i)})\}_{1 \leq i \leq n}$ be an i.i.d sample of (X, Y) , we consider for any $j \in \{1, \dots, q\}$ and any $h \in \{1, \dots, r_j\}$:

$$\hat{n}_{jh} = \sum_{i=1}^n \mathbf{1}_{\{Y_j^{(i)} \in I_h^{(j)}\}}, \quad \hat{p}_{jh}^n = \frac{\hat{n}_{jh}}{n}, \quad \bar{X}_{jh}^n = \frac{1}{\hat{n}_{jh}} \sum_{i=1}^n \mathbf{1}_{\{Y_j^{(i)} \in I_h^{(j)}\}} X^{(i)}, \quad \bar{X}^n = \frac{1}{n} \sum_{i=1}^n X^{(i)}, \tag{1}$$

where $Y_j^{(i)}$ is the j -th coordinate of $Y^{(i)}$, and $\mathbf{1}_A$ denotes the indicator function of A . Then, we estimate Λ_j by the random operator

$$\hat{\Lambda}_j^n = \sum_{h=1}^{r_j} \hat{p}_{jh}^n (\bar{X}_{jh}^n - \bar{X}^n) \otimes (\bar{X}_{jh}^n - \bar{X}^n),$$

and, putting $\widehat{\Lambda}^n = \sum_{j=1}^q \widehat{\Lambda}_j^n$, we take as test statistic the random variable

$$\widehat{S}^{(n)} = tr(\widehat{\Lambda}^n).$$

It is a strongly consistent estimator of $tr(\widetilde{\Lambda})$. Indeed, the strong law of large numbers ensures the almost sure convergence of \widehat{p}_{jh}^n (resp. \overline{X}^n ; resp. \overline{X}_{jh}^n) to p_{jh} (resp. μ ; resp. μ_{jh}) as $n \rightarrow +\infty$ and, therefore, that of $\widehat{\tau}_{jh}^n = \overline{X}_{jh}^n - \overline{X}^n$ to τ_{jh} . As the bilinear map $(x, y) \in \mathbb{R}^p \times \mathbb{R}^p \mapsto x \otimes y \in \mathcal{L}(\mathbb{R}^p)$ is continuous, we then deduce the almost sure convergence of $\widehat{\Lambda}_j^n$ to $\widetilde{\Lambda}_j$ as $n \rightarrow +\infty$, which implies that of $\widehat{S}^{(n)}$ to $tr(\widetilde{\Lambda})$.

Now, we will give the asymptotic distribution of $\widehat{S}^{(n)}$ under \mathcal{H}_0 . For $j \in \{1, \dots, q\}$, let us introduce the diagonal matrix $\Delta_j = diag(p_{j1}, p_{j2}, \dots, p_{jr_j})$ and put $\Gamma_j = \Delta_j^{-1} \otimes^K I_p$, where \otimes^K denotes the Kronecker product and I_p is the $p \times p$ identity matrix. Then, we consider the matrices

$$\Gamma = \begin{pmatrix} \Gamma_1 & 0 & \dots & 0 \\ 0 & \Gamma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Gamma_q \end{pmatrix} \tag{2}$$

and

$$\Sigma = \begin{pmatrix} \sigma_{1111} & \sigma_{1112} & \dots & \sigma_{11qr_q} \\ \sigma_{1211} & \sigma_{1212} & \dots & \sigma_{12qr_q} \\ \vdots & \vdots & \dots & \vdots \\ \sigma_{qr_q11} & \sigma_{qr_q12} & \dots & \sigma_{qr_qqr_q} \end{pmatrix}, \tag{3}$$

where

$$\sigma_{ikj\ell} = V^{(ikj\ell)} + p_{ik}p_{j\ell} (V - V^{(ik)} - V^{(j\ell)}),$$

with

$$\begin{aligned} V^{(ik)} &= \mathbb{E}((X - \mu_{ik}) \otimes (X - \mu_{ik}) | Y_i \in I_k), \\ V^{(ikj\ell)} &= \mathbb{E}((\mathbf{1}_{\{Y_i \in I_k\}} (X - \mu_{ik})) \otimes (\mathbf{1}_{\{Y_j \in I_\ell\}} (X - \mu_{j\ell}))), \end{aligned}$$

and

$$V = \mathbb{E}((X - \mu) \otimes (X - \mu)).$$

Then, we have:

Theorem 1. Under \mathcal{H}_0 , $n\widehat{S}^{(n)}$ converges in distribution, as $n \rightarrow +\infty$, to $\mathcal{Q} = U'\Gamma U$ where U is a centered random vector having a normal distribution in \mathbb{R}^{pq} with covariance operator equal to Σ .

2.3. The test procedure

For a given significance level $\alpha \in]0, 1[$, the hypothesis \mathcal{H}_0 will be rejected if $\mathbb{F}_{\mathcal{Q}}\left(n\widehat{S}^{(n)}\right) > 1 - \alpha$, where $\mathbb{F}_{\mathcal{Q}}$ denotes the cumulative distribution function of \mathcal{Q} . Since \mathcal{Q} is a quadratic form of a normally distributed random vector, $\mathbb{F}_{\mathcal{Q}}$ can be computed or approximated by using formulas given in [Mathai and Provost, \(1992\)](#) and which involve the eigenvalues of $\Sigma^{\frac{1}{2}}\Gamma\Sigma^{\frac{1}{2}}$. In practice Σ and Γ are unknown. So, they are to be replaced by consistent estimators. For estimating Γ , we consider

$$\widehat{\Gamma}^{(n)} = \begin{pmatrix} \widehat{\Gamma}_1^n & 0 & \dots & 0 \\ 0 & \widehat{\Gamma}_2^n & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \widehat{\Gamma}_q^n \end{pmatrix} \quad (4)$$

where $\widehat{\Gamma}_j^n = \left(\widehat{\Delta}_j^n\right)^{-1} \otimes^K I_p$, with $\widehat{\Delta}_j^n = \text{diag}\left(\widehat{p}_{j1}^n, \widehat{p}_{j2}^n, \dots, \widehat{p}_{jr_j}^n\right)$. The estimation of Σ is achieved by considering the $pq \times pq$ matrix $\widehat{\Sigma}_n$ given by

$$\widehat{\Sigma}_n = \begin{pmatrix} \widehat{\sigma}_{1111}^n & \widehat{\sigma}_{1112}^n & \dots & \widehat{\sigma}_{11qr_q}^n \\ \widehat{\sigma}_{1211}^n & \widehat{\sigma}_{1212}^n & \dots & \widehat{\sigma}_{12qr_q}^n \\ \vdots & \vdots & \vdots & \vdots \\ \widehat{\sigma}_{qr_q11}^n & \widehat{\sigma}_{qr_q12}^n & \dots & \widehat{\sigma}_{qr_qqr_q}^n \end{pmatrix}, \quad (5)$$

where

$$\widehat{\sigma}_{ikj\ell}^n = V_n^{(ikj\ell)} + \widehat{p}_{ik}^n \widehat{p}_{j\ell}^n \left(\widehat{V}_n - \widehat{V}_n^{(ik)} - \widehat{V}_n^{(j\ell)}\right)$$

with

$$\begin{aligned} \widehat{V}_n^{(ikj\ell)} &= \frac{1}{n} \sum_{t=1}^n \left(\mathbf{1}_{\{Y_i^{(t)} \in I_k\}} \left(X^{(t)} - \overline{X}_{ik}^n\right)\right) \otimes \left(\mathbf{1}_{\{Y_j^{(t)} \in I_\ell\}} \left(X^{(t)} - \overline{X}_{j\ell}^n\right)\right) \\ \widehat{V}_n^{(jk)} &= \frac{1}{n_{jk}} \sum_{i=1}^n \mathbf{1}_{\{Y_j^{(i)} \in I_k\}} \left(X^{(i)} - \overline{X}_{jk}^n\right) \otimes \left(X^{(i)} - \overline{X}_{jk}^n\right) \\ \widehat{V}_n &= \frac{1}{n} \sum_{i=1}^n \left(X^{(i)} - \overline{X}^n\right) \otimes \left(X^{(i)} - \overline{X}^n\right). \end{aligned}$$

Remark 2. Practical implementation can be done by using Remark 1 that shows that, by identifying each vector with its matrix relative to canonical basis, we can write :

$$\widehat{\Lambda}_j^n = \sum_{h=1}^{r_j} \widehat{p}_{jh}^n \widehat{\tau}_{jh}^n (\widehat{\tau}_{jh}^n)' \tag{6}$$

$$\widehat{V}_n^{(ikj\ell)} = \frac{1}{n} \sum_{t=1}^n \left(\mathbf{1}_{\{Y_j^{(t)} \in I_\ell\}} \left(X^{(t)} - \overline{X}_{j\ell}^n \right) \right) \left(\mathbf{1}_{\{Y_i^{(t)} \in I_k\}} \left(X^{(t)} - \overline{X}_{ik}^n \right) \right)' \tag{7}$$

$$\widehat{V}_n^{(jk)} = \frac{1}{n_{jk}} \sum_{i=1}^n \mathbf{1}_{\{Y_j^{(i)} \in I_k\}} \left(X^{(i)} - \overline{X}_{jk}^n \right) \left(X^{(i)} - \overline{X}_{jk}^n \right)' \tag{8}$$

$$\widehat{V}_n = \frac{1}{n} \sum_{i=1}^n \left(X^{(i)} - \overline{X}^n \right) \left(X^{(i)} - \overline{X}^n \right)' \tag{9}$$

Remark 3. The proposed method is achieved from the following algorithm:

1. Compute \overline{X}^n and for $j = 1, \dots, q$ and $h = 1, \dots, r_j$ compute $\widehat{n}_{jh}^n, \overline{X}_{jh}^n$ and \widehat{p}_{jh}^n from (1). Then compute $\widehat{\tau}_{jh}^n = \overline{X}_{jh}^n - \overline{X}^n$ and $\widehat{\Lambda}_j^n$ from (6) .
2. Compute $\widehat{S}^{(n)} = tr \left(\sum_{j=1}^q \widehat{\Lambda}_j^n \right)$ and, for $j = 1, \dots, q$, compute $\widehat{\Delta}_j^n, \widehat{\Gamma}_j^n = \left(\widehat{\Delta}_j^n \right)^{-1} \otimes^K I_p$ and $\widehat{\Gamma}^{(n)}$ by using (4) .
3. Compute \widehat{V}_n from (9) and for $i, j = 1, \dots, q, k = 1, \dots, r_i$ and $\ell = 1, \dots, r_j$, compute $\widehat{V}_n^{(ikj\ell)}$ and $\widehat{V}_n^{(jk)}$ from (7) and (8) respectively.
4. For $i, j = 1, \dots, q, k = 1, \dots, r_i$ and $\ell = 1, \dots, r_j$, compute

$$\widehat{\sigma}_{ikj\ell}^n = \widehat{V}_n^{(ikj\ell)} + p_{ik}^n p_{j\ell}^n \left(\widehat{V}_n - \widehat{V}_n^{(ik)} - \widehat{V}_n^{(j\ell)} \right) .$$

5. Compute the matrix $\widehat{\Sigma}^{(n)}$ given in (5).
6. Compute the eigenvalues $\lambda_1, \dots, \lambda_{qr_q}$ of $\left(\widehat{\Sigma}^{(n)} \right)^{1/2} \widehat{\Gamma}^{(n)} \left(\widehat{\Sigma}^{(n)} \right)^{1/2}$. Then consider the function $\mathbb{F}_{\mathcal{Q}}(t) := P \left(\chi_{\nu}^2 < t/c \right)$ where

$$c = \frac{\left(\sum_{j=1}^{qr_q} \lambda_j^2 \right)}{\left(\sum_{j=1}^{qr_q} \lambda_j \right)} \text{ and } \nu = \frac{\left(\sum_{j=1}^{qr_q} \lambda_j \right)^2}{\left(\sum_{j=1}^{qr_q} \lambda_j^2 \right)} .$$

7. Compute $p_c = 1 - \mathbb{F}_{\mathcal{Q}} \left(n \widehat{S}^{(n)} \right)$. If $p_c \geq \alpha$ then accept \mathcal{H}_0 ; otherwise reject it.

3. Simulation results

In order to check the efficacy of the proposed method and to compare it with that of existing methods, a simulation study is performed here. We computed empirical sizes and empirical powers over 1000 replications, with nominal significance level $\alpha = 0.05, 0.10$, from our

method, denoted here as REG, and three known methods. These methods are the likelihood ratio test (LRT), the method based on ranks given in Cl eroux, (1995) and denoted here as CLL, and the method introduced in Meintanis and Iliopoulos, (2008) denoted by MI. For sample sizes $n = 50, 100, 200, 300, 400, 500$, we generated 1000 independent replicates of a pair $Z = (X', Y)'$ of random vectors by using the following models:

Model 1: Z has a centered normal distribution in \mathbb{R}^{10} with covariance matrix C given by

$$C = \begin{pmatrix} I_5 & \gamma I_5 \\ \gamma I_5 & I_5 \end{pmatrix}$$

where I_5 denotes the 5×5 identity matrix and γ is a real belonging to $[0, 1]$.

Model 2: Z is a two-dimensional random vector with coordinates defined by

$$X = \sqrt{\frac{3}{20}}(aX_1 + bY_1) \text{ and } Y = \sqrt{\frac{3}{20}}(bX_1 + aY_1),$$

where X_1 and Y_1 are independent random variables having a student distribution with 5 degrees of freedom, and

$$a = \sqrt{1 + \gamma} + \sqrt{1 - \gamma}, \quad b = \sqrt{1 + \gamma} - \sqrt{1 - \gamma},$$

where γ belongs to $[0, 1]$.

For both models, \mathcal{H}_0 holds if and only if $\gamma = 0$. Table 1 gives outputs for empirical sizes from our method. Satisfactory results are provided, except for low sample size ($n = 50$). The most accurate results are obtained for $n \geq 200$. The obtained results for empirical powers, for γ from 0 up to 0.4, are given in Figures 1 to 3 for Model 1, and in Figures 4 to 6 for Model 2. They show that, for moderate values of γ and sufficiently large values of n , our method outperforms all the considered existing methods except LRT that gives the best results in all the tackled situations. However, for $n = 100$ our method is also outperformed by CLL, and by MI when γ is small. For large values of γ all the methods give the same empirical power equal to 1.

n	Model 1		Model 2	
	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.10$
50	0.018	0.052	0.024	0.06
100	0.027	0.078	0.038	0.084
200	0.047	0.089	0.032	0.093
300	0.047	0.104	0.047	0.091
400	0.038	0.083	0.042	0.094
500	0.044	0.097	0.046	0.095

Table 1. Empirical sizes from REG, with sample size n and nominal significance level α .

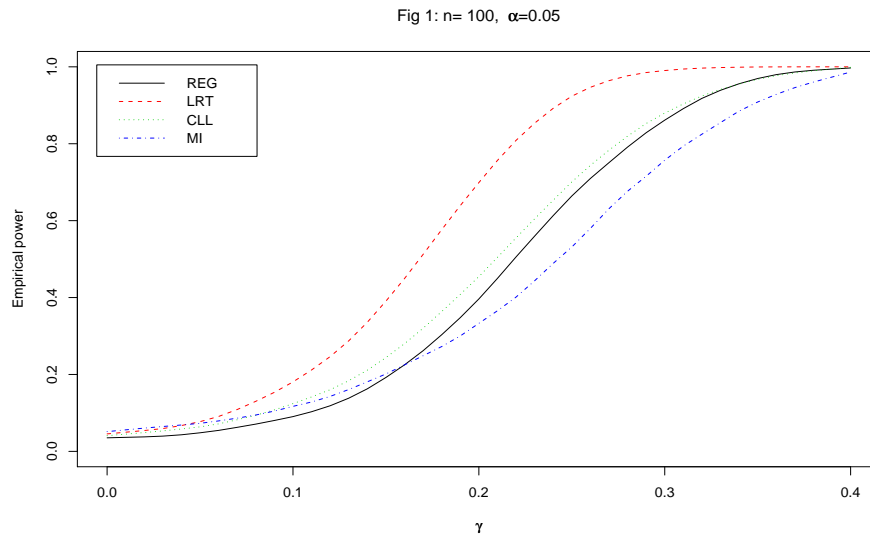


Fig. 1. Empirical power versus γ for Model 1, $n = 100$ and $\alpha = 0.05$.

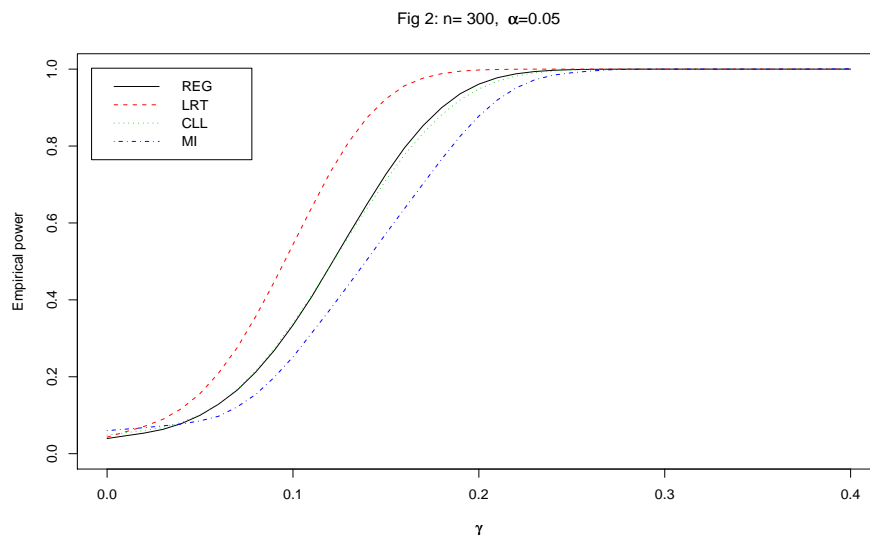


Fig. 2. Empirical power versus γ for Model 1, $n = 300$ and $\alpha = 0.05$.

4. Proofs

4.1. Lemmas

Putting $E = \mathbb{R}^p$ and $F = \mathbb{R}^{r_1} \times \mathbb{R}^{r_2} \times \dots \times \mathbb{R}^{r_q}$, we introduce the operators

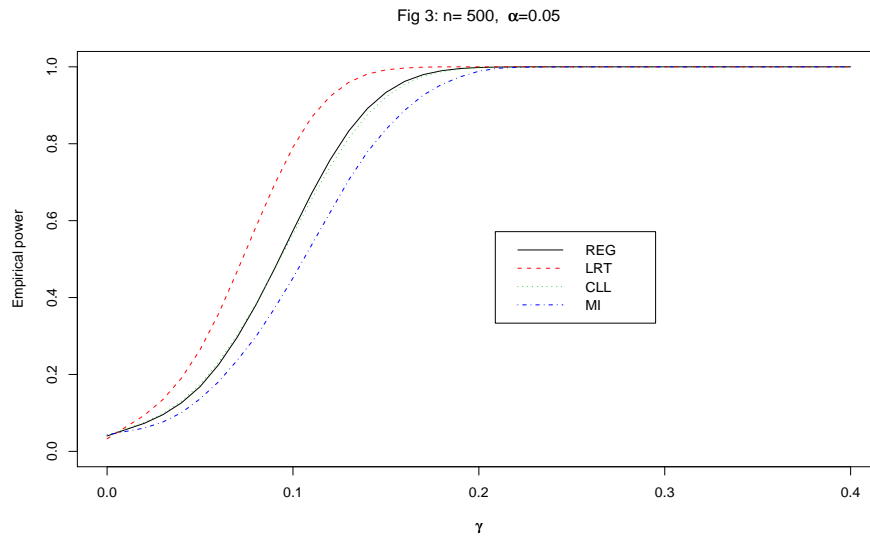


Fig. 3. Empirical power versus γ for Model 1, $n = 500$ and $\alpha = 0.05$.

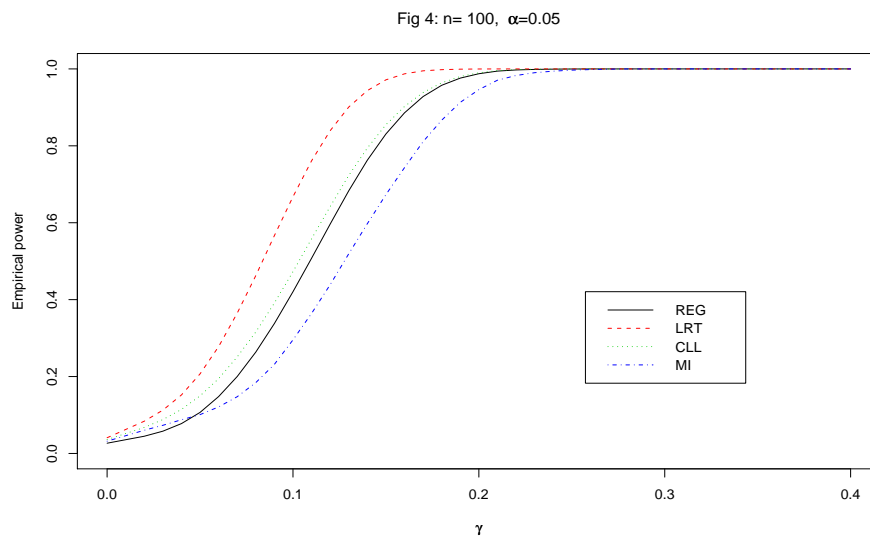


Fig. 4. Empirical power versus γ for Model 2, $n = 100$ and $\alpha = 0.05$.

$$\Pi_0 : \begin{pmatrix} u \\ v \end{pmatrix} \in E \times F \mapsto \begin{pmatrix} u \\ 0 \end{pmatrix} \in E \times F,$$

$$\Pi_1 : \begin{pmatrix} A \\ x \end{pmatrix} \in \mathcal{L}(E \times F) \times (E \times F) \mapsto A \in \mathcal{L}(E \times F),$$

$$\Pi_2 : \begin{pmatrix} A \\ x \end{pmatrix} \in \mathcal{L}(E \times F) \times (E \times F) \mapsto x \in E \times F.$$

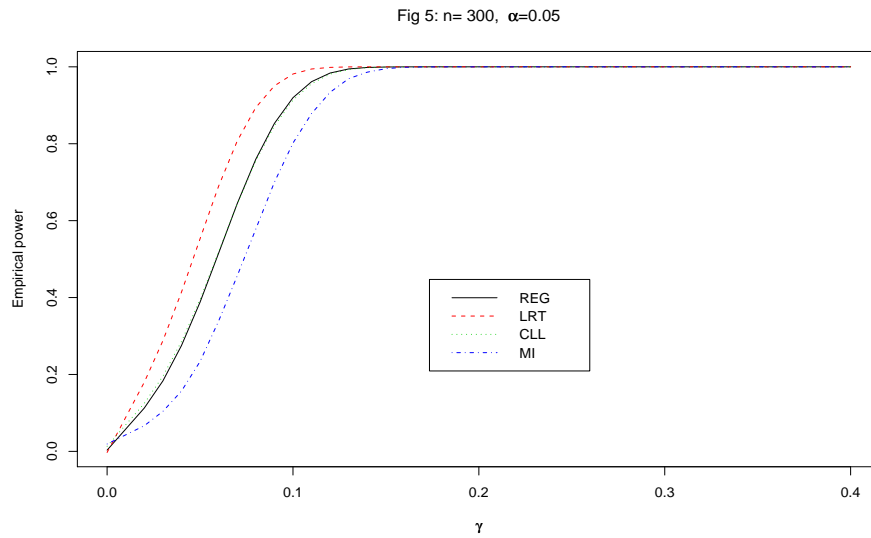


Fig. 5. Empirical power versus γ for Model 2, $n = 300$ and $\alpha = 0.05$.

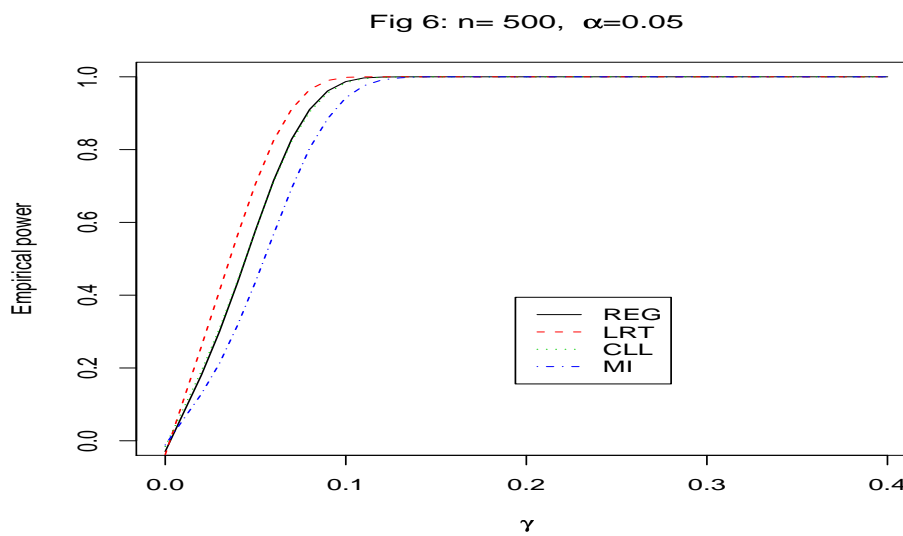


Fig. 6. Empirical power versus γ for Model 2, $n = 500$ and $\alpha = 0.05$.

Furthermore, we consider the random vectors

$$W_j := \sum_{h=1}^{r_j} \mathbf{1}_{\{Y_j \in I_h^{(j)}\}} f_h^{(j)}, W_j^{(i)} := \sum_{h=1}^{r_j} \mathbf{1}_{\{Y_j^{(i)} \in I_h^{(j)}\}} f_h^{(j)}$$

where $\{f_1^{(j)}, \dots, f_{r_j}^{(j)}\}$ is the canonical basis of \mathbb{R}^{r_j} , and the random vectors valued into $E \times F$ given by:

$$Z = \begin{pmatrix} X - \mu \\ W_1 \\ \vdots \\ W_q \end{pmatrix}, Z^{(i)} = \begin{pmatrix} X^{(i)} - \bar{X}^n \\ W_1^{(i)} \\ \vdots \\ W_q^{(i)} \end{pmatrix}, Z_0 = \begin{pmatrix} X \\ W_1 \\ \vdots \\ W_q \end{pmatrix}.$$

Then, putting

$$T = \begin{pmatrix} Z_0 \otimes Z_0 \\ Z_0 \end{pmatrix}, V_Z = \mathbb{E}(Z \otimes Z), V_Z^{(n)} = \frac{1}{n} \sum_{i=1}^n Z^{(i)} \otimes Z^{(i)},$$

and considering the linear map φ from $\mathcal{L}(E \times F) \times (E \times F)$ to $\mathcal{L}(E \times F)$ defined by:

$$\begin{aligned} \varphi(S) &= \Pi_1(S) - \Pi_2(S) \otimes \Pi_0(\mu_1) - \mu_1 \otimes \Pi_0(\Pi_2(S)) - \Pi_0(\Pi_2(S)) \otimes \mu_1 \\ &\quad - \Pi_0(\mu_1) \otimes (\Pi_2(S)) + \Pi_0(\Pi_2(S)) \otimes \Pi_0(\mu_1) + \Pi_0(\mu_1) \otimes \Pi_0(\Pi_2(S)) \end{aligned}$$

where $\mu_1 = \mathbb{E}(Z_0)$, and denoting by $\tilde{\otimes}$ the tensor product between elements of $\mathcal{L}(E \times F)$ (associated with the inner product of this space defined by: $\langle A, B \rangle = \text{tr}(AB^*)$), we have:

Lemma 1. *As $n \rightarrow +\infty$, $H_n = \sqrt{n} \left(V_Z^{(n)} - V_Z \right)$ converges in distribution to a random variable H having a centered normal distribution in $L(E \times F)$ with covariance operator equal to $E[(\varphi(T - \mathbb{E}(T))) \tilde{\otimes} (\varphi(T - \mathbb{E}(T)))]$.*

Proof.

Clearly, $Z = Z_0 - \Pi_0(\mu_1)$ and $Z^{(i)} = Z_0^{(i)} - \Pi_0(\bar{Z}_0^n)$, where

$$Z_0^{(i)} = \begin{pmatrix} X^{(i)} \\ W_1^{(i)} \\ \vdots \\ W_q^{(i)} \end{pmatrix} \quad \text{and} \quad \bar{Z}_0^n = \frac{1}{n} \sum_{i=1}^n Z_0^{(i)}.$$

Therefore,

$$\begin{aligned} H_n &= \sqrt{n} \left[\frac{1}{n} \sum_{i=1}^n Z_0^{(i)} \otimes Z_0^{(i)} - \mathbb{E}(Z_0 \otimes Z_0) - (\bar{Z}_0^n - \mu_1) \otimes \Pi_0(\bar{Z}_0^n) \right. \\ &\quad - \mu_1 \otimes \Pi_0(\bar{Z}_0^n - \mu_1) - \Pi_0(\bar{Z}_0^n - \mu_1) \otimes \bar{Z}_0^n - \Pi_0(\mu_1) \otimes (\bar{Z}_0^n - \mu_1) \\ &\quad \left. + \Pi_0(\bar{Z}_0^n - \mu_1) \otimes \Pi_0(\bar{Z}_0^n) + \Pi_0(\mu_1) \otimes \Pi_0(\bar{Z}_0^n - \mu_1) \right]. \end{aligned}$$

Considering the $\mathcal{L}(E \times F) \times (E \times F)$ -valued random variables

$$T^{(i)} = \begin{pmatrix} Z_0^{(i)} \otimes Z_0^{(i)} \\ Z_0^{(i)} \end{pmatrix}, \quad i = 1, 2, \dots, n,$$

and putting $K_n = \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n T^{(i)} - \mathbb{E}(T) \right)$, we can write

$$\begin{aligned} H_n &= \Pi_1(K_n) - \Pi_2(K_n) \otimes \Pi_0(\bar{Z}_0^n) - \mu_1 \otimes \Pi_0(\Pi_2(K_n)) \\ &\quad - \Pi_0(\Pi_2(K_n)) \otimes \bar{Z}_0^n - \Pi_0(\mu_1) \otimes (\Pi_2(K_n)) \\ &\quad + \Pi_0(\Pi_2(K_n)) \otimes \Pi_0(\bar{Z}_0^n) + \Pi_0(\mu_1) \otimes \Pi_0(\Pi_2(K_n)) \\ &= \varphi_n(K_n), \end{aligned} \tag{10}$$

where φ_n is the random operator from $\mathcal{L}(E \times F) \times (E \times F)$ to $\mathcal{L}(E \times F)$ defined as:

$$\begin{aligned} \varphi_n(S) &= \Pi_1(S) - \Pi_2(S) \otimes \Pi_0(\bar{Z}_0^n) - \mu_1 \otimes \Pi_0(\Pi_2(S)) - \Pi_0(\Pi_2(S)) \otimes \bar{Z}_0^n \\ &\quad - \Pi_0(\mu_1) \otimes (\Pi_2(S)) + \Pi_0(\Pi_2(S)) \otimes \Pi_0(\bar{Z}_0^n) + \Pi_0(\mu_1) \otimes \Pi_0(\Pi_2(S)) \\ &= g(S, \bar{Z}_0^n), \end{aligned} \tag{11}$$

and g is the continuous map from $(\mathcal{L}(E \times F) \times (E \times F)) \times (E \times F)$ to $\mathcal{L}(E \times F)$ defined as:

$$\begin{aligned} g(S, C) &= \Pi_1(S) - \Pi_2(S) \otimes \Pi_0(C) - \mu_1 \otimes \Pi_0(\Pi_2(S)) - \Pi_0(\Pi_2(S)) \otimes C \\ &\quad - \Pi_0(\mu_1) \otimes (\Pi_2(S)) + \Pi_0(\Pi_2(S)) \otimes \Pi_0(C) + \Pi_0(\mu_1) \otimes \Pi_0(\Pi_2(S)). \end{aligned}$$

The central limit theorem in $\mathcal{L}(E \times F) \times (E \times F)$ ensures that K_n converges in distribution, as $n \rightarrow +\infty$, to a random operator K having a centered normal distribution in $\mathcal{L}(E \times F) \times (E \times F)$ with covariance operator $\Gamma_1 = \mathbb{E}((T - \mathbb{E}(T)) \otimes (T - \mathbb{E}(T)))$. Furthermore, from the strong law of large numbers, we have the almost sure convergence, in $E \times F$, of \bar{Z}_0^n to μ_1 as $n \rightarrow +\infty$. We deduce (see Billingsley, (1968)) that (K_n, \bar{Z}_0^n) converges in distribution to (K, μ_1) , as $n \rightarrow +\infty$. From the continuity of g , it follows that $g(K_n, \bar{Z}_0^n)$ converges in distribution, as $n \rightarrow +\infty$, to $g(K, \mu_1) = \varphi(K)$. Then, from (10) and (11), H_n converges in distribution to $H = \varphi(K)$. Since φ is linear, H has a centered normal distribution in $\mathcal{L}(E \times F)$ with covariance operator $\Gamma_2 = \mathbb{E}((\varphi(T - \mathbb{E}(T))) \otimes (\varphi(T - \mathbb{E}(T))))$. \square

The following lemma gives the limiting distribution of $\hat{\Lambda}^n$ under \mathcal{H}_0 . Any operator S of $\mathcal{L}(E \times F)$ can be partitioned in the form

$$S = \begin{pmatrix} S_{00} & S_{01} & \dots & S_{0q} \\ S_{10} & S_{11} & \dots & S_{1q} \\ \vdots & \vdots & \vdots & \vdots \\ S_{q0} & S_{q1} & \dots & S_{qq} \end{pmatrix}$$

where

$$S_{00} \in \mathcal{L}(\mathbb{R}^p), S_{0j} \in \mathcal{L}(\mathbb{R}^{r_j}, \mathbb{R}^p), S_{j0} \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^{r_j}), S_{kl} \in \mathcal{L}(\mathbb{R}^{r_l}, \mathbb{R}^{r_k})$$

with $1 \leq j, k, l \leq q$. In this context, let us consider the operators

$$\Pi_{0j} : S \in \mathcal{L}(E \times F) \mapsto S_{0j} \in \mathcal{L}(\mathbb{R}^{r_j}, \mathbb{R}^p)$$

and the map

$$\psi : A \in \mathcal{L}(E \times F) \mapsto \sum_{j=1}^q \sum_{h=1}^{r_j} p_{jh}^{-1} \left(\Pi_{0j}(A) f_h^{(j)} \right) \otimes \left(\Pi_{0j}(A) f_h^{(j)} \right) \in \mathcal{L}(E).$$

Then, we have:

Lemma 2. Under \mathcal{H}_0 , $n\widehat{\Lambda}^n$ converges in distribution to $\psi(H)$ as $n \rightarrow +\infty$.

Proof.

Since we have $\tau_{jh} = 0$ under \mathcal{H}_0 , we can write

$$n\widehat{\Lambda}^n = \sum_{j=1}^q \sum_{h=1}^{r_j} (\widehat{p}_{jh}^n)^{-1} \sqrt{n} (\widehat{p}_{jh}^n \widehat{\tau}_{jh}^n - p_{jh} \tau_{jh}) \otimes \sqrt{n} (\widehat{p}_{jh}^n \widehat{\tau}_{jh}^n - p_{jh} \tau_{jh}).$$

Moreover, from

$$\begin{aligned} \mathbb{E}(W_j \otimes (X - \mu)) &= \mathbb{E} \left(\left(\sum_{h=1}^{r_j} \mathbf{1}_{\{Y_j \in I_h^{(j)}\}} f_h^{(j)} \right) \otimes X - \left(\sum_{h=1}^{r_j} \mathbf{1}_{\{Y_j \in I_h^{(j)}\}} f_h^{(j)} \right) \otimes \mu \right) \\ &= \sum_{h=1}^{r_j} f_h^{(j)} \otimes (p_{jh} \tau_{jh}) \end{aligned}$$

and $n^{-1} \sum_{i=1}^n (W_j^{(i)} \otimes (X^{(i)} - \bar{X}^n)) = \sum_{h=1}^{r_j} f_h^{(j)} \otimes (\widehat{p}_{jh}^n \widehat{\tau}_{jh}^n)$, we easily obtain the equality

$\Pi_{0j}(H_n) = \sum_{h=1}^{r_j} f_h^{(j)} \otimes (\sqrt{n} (p_{jh} \tau_{jh} - \widehat{p}_{jh}^n \widehat{\tau}_{jh}^n))$. Thus:

$$n\widehat{\Lambda}^n = \sum_{j=1}^q \sum_{h=1}^{r_j} (\widehat{p}_{jh}^n)^{-1} \left(\Pi_{0j}(H_n) f_h^{(j)} \right) \otimes \left(\Pi_{0j}(H_n) f_h^{(j)} \right) = \psi_n(H_n),$$

where ψ_n is the random map from $\mathcal{L}(E \times F)$ to $\mathcal{L}(E)$ defined as:

$$\psi_n(S) = \sum_{j=1}^q \sum_{h=1}^{r_j} (\widehat{p}_{jh}^n)^{-1} \left(\Pi_{0j}(S) f_h^{(j)} \right) \otimes \left(\Pi_{0j}(S) f_h^{(j)} \right).$$

Putting $\Delta_n = \psi_n(H_n) - \psi(H_n)$, we have

$$\Delta_n = \sum_{j=1}^q \sum_{h=1}^{r_j} \left(\Pi_{0j}(H_n) f_h^{(j)} \right) \otimes \left((\widehat{p}_{jh}^n)^{-1} - p_{jh}^{-1} \right) \Pi_{0j}(H_n) f_h^{(j)} = \sum_{j=1}^q \sum_{h=1}^{r_j} g_{jh}(H_n, \widehat{p}_{jh}^n),$$

where

$$g_{jh} : (A, x) \in \mathcal{L}(E \times F) \times \mathbb{R}_+^* \mapsto \left(\Pi_{0j}(A) f_h^{(j)} \right) \otimes \left((x^{-1} - p_{jh}^{-1}) \Pi_{0j}(A) f_h^{(j)} \right).$$

Since \widehat{p}_{jh}^n converges in probability to p_{jh} as $n \rightarrow +\infty$, we deduce from Lemma 1 and the continuity of g_{jh} that $g_{jh}(H_n, \widehat{p}_{jh}^n)$ converges in distribution to $g_{jh}(H, p_{jh})$, as $n \rightarrow +\infty$. Clearly, $g_{jh}(H, p_{jh}) = 0$; then, the preceding convergence property is a convergence in probability. Consequently, Δ_n converges in probability to 0 as $n \rightarrow +\infty$ and, therefore, $\psi_n(H_n)$ and $\psi(H_n)$ have the same limiting distribution. As ψ is continuous and H_n converges in distribution to H , $n\widehat{\Lambda}^n$ converges in distribution to $\psi(H)$. \square

4.2. Proof of Theorem 1

From Lemma 2, we deduce that under \mathcal{H}_0 , $ntr(\widehat{\Lambda}^n)$ converges in distribution, as $n \rightarrow +\infty$, to $\mathcal{Q} = tr(\psi(H))$. Now, it remains to prove that \mathcal{Q} has the required expression. We have

$$\begin{aligned} tr(\psi(H)) &= \sum_{j=1}^q \sum_{h=1}^{r_j} p_{jh}^{-1} tr \left(\left(\Pi_{0j}(H) f_h^{(j)} \right) \otimes \left(\Pi_{0j}(H) f_h^{(j)} \right) \right) \\ &= \sum_{j=1}^q \sum_{h=1}^{r_j} p_{jh}^{-1} \left\langle \Pi_{0j}(H) f_h^{(j)}, \left(\Pi_{0j}(H) f_h^{(j)} \right) \right\rangle \\ &= \sum_{j=1}^q \sum_{h=1}^{r_j} \left(\Pi_{0j}(H) f_h^{(j)} \right)' \left(p_{jh}^{-1} I_p \right) \left(\Pi_{0j}(H) f_h^{(j)} \right) \\ &= \sum_{j=1}^q \mathcal{U}_j' \Gamma_j \mathcal{U}_j, \end{aligned}$$

where

$$\mathcal{U}_j = \begin{pmatrix} \Pi_{0j}(H) f_1^{(j)} \\ \Pi_{0j}(H) f_2^{(j)} \\ \vdots \\ \Pi_{0j}(H) f_{r_j}^{(j)} \end{pmatrix} = \Phi_j(H)$$

and

$$\Gamma_j = \begin{pmatrix} p_{j1}^{-1} I_p & 0 & \dots & 0 \\ 0 & p_{j2}^{-1} I_p & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & p_{jr_j}^{-1} I_p \end{pmatrix} = \Delta_j^{-1} \otimes^K I_p$$

with $\Delta_j = \text{diag}(p_{j1}, p_{j2}, \dots, p_{jr_j})$. Putting

$$\mathcal{U} = \begin{pmatrix} \mathcal{U}_1 \\ \mathcal{U}_2 \\ \vdots \\ \mathcal{U}_q \end{pmatrix} = \begin{pmatrix} \Phi_1(H) \\ \Phi_2(H) \\ \vdots \\ \Phi_q(H) \end{pmatrix} = \Phi(H) \tag{12}$$

and taking Γ as defined in (2), we clearly have $\mathcal{U}'\Gamma\mathcal{U} = \mathcal{Q}$. Since Φ is linear and H is centered and normally distributed with covariance operator equal to that of $\varphi(T)$, we deduce from (12) that \mathcal{U} is also centered and normally distributed; its covariance operator Σ equals that of $\Phi(\varphi(T))$, that is

$$\Sigma = \mathbb{E}[(\Phi(\varphi(T))) \otimes (\Phi(\varphi(T)))] - \mathbb{E}(\Phi(\varphi(T))) \otimes \mathbb{E}(\Phi(\varphi(T))), \tag{13}$$

where

$$\begin{aligned} \Phi(\varphi(T)) &= \Phi(Z_0 \otimes Z_0) - \Phi(Z_0 \otimes \Pi_0(\mu_1)) - \Phi(\mu_1 \otimes \Pi_0(Z_0)) \\ &\quad - \Phi(\Pi_0(Z_0) \otimes \mu_1) - \Phi(\Pi_0(\mu_1) \otimes Z_0) \\ &\quad + \Phi(\Pi_0(Z_0) \otimes \Pi_0(\mu_1)) + \Phi(\Pi_0(\mu_1) \otimes \Pi_0(Z_0)). \end{aligned}$$

Since

$$\begin{aligned} \Pi_{0j}(Z_0 \otimes Z_0) f_h^{(j)} &= (W_j \otimes X) f_h^{(j)} = \mathbf{1}_{\{Y_j \in I_h^{(j)}\}} X; \\ \Pi_{0j}(Z_0 \otimes \Pi_0(\mu)) f_h^{(j)} &= (W_j \otimes \mu) f_h^{(j)} = \mathbf{1}_{\{Y_j \in I_h^{(j)}\}} \mu; \\ \Pi_{0j}(\mu \otimes \Pi_0(Z_0)) f_h^{(j)} &= (\mathbb{E}(W_j) \otimes X) f_h^{(j)} = p_{jh} X; \\ \Pi_{0j}(\Pi_0(Z_0) \otimes \mu) &= \Pi_{0j}(\Pi_0(\mu) \otimes Z_0) = 0; \\ \Pi_{0j}(\Pi_0(Z_0) \otimes \Pi_0(\mu)) &= \Pi_{0j}(\Pi_0(\mu) \otimes \Pi_0(Z_0)) = 0; \end{aligned}$$

we obtain $\Phi(\varphi(T)) = (u_{11}, \dots, u_{1r_1}, \dots, u_{q1}, \dots, u_{qr_q})'$, where

$$u_{jh} = \mathbf{1}_{\{Y_j \in I_h^{(j)}\}} (X - \mu) - p_{jh} X.$$

Thus, we deduce from (13) that Σ has the form given in (3) with $\sigma_{ikj\ell} = \mathbb{E}(u_{ik} \otimes u_{j\ell}) - \mathbb{E}(u_{ik}) \otimes \mathbb{E}(u_{j\ell})$, where

$$\begin{aligned} \mathbb{E}(u_{ik} \otimes u_{j\ell}) &= \mathbb{E}(\mathbf{1}_{\{Y_i \in I_k\}} (X - \mu_{ik}) \otimes \mathbf{1}_{\{Y_j \in I_\ell\}} (X - \mu_{j\ell})) + p_{ik} p_{j\ell} \mathbb{E}(X \otimes X) \\ &\quad - p_{j\ell} p_{ik} \mathbb{E}((X - \mu_{ik}) \otimes (X - \mu_{ik}) | Y_i \in I_k) \\ &\quad - p_{j\ell} p_{ik} \mathbb{E}((X - \mu_{j\ell}) \otimes (X - \mu_{j\ell}) | Y_j \in I_\ell) \\ &\quad - p_{j\ell} \mathbb{E}(\mathbf{1}_{\{Y_i \in I_k\}} (X - \mu_{ik}) \otimes \mu_{ik}) - p_{ik} \mathbb{E}(\mu_{j\ell} \otimes \mathbf{1}_{\{Y_j \in I_\ell\}} (X - \mu_{j\ell})) \end{aligned}$$

and

$$\mathbb{E}(u_{ik}) = \mathbb{E}\left(\mathbf{1}_{\{Y_i \in I_k^{(i)}\}} (X - \mu) - p_{ik} X\right) = p_{ik} \mu_{ik} - p_{ik} \mu - p_{ik} \mu = -p_{ik} \mu.$$

Since, under \mathcal{H}_0 , we have $\mu_{ik} = \mu$, it follows that

$$\begin{aligned}\mathbb{E}(\mathbf{1}_{\{Y_i \in I_k\}}(X - \mu_{ik}) \otimes \mu_{ik}) &= (\mathbb{E}(\mathbf{1}_{\{Y_i \in I_k\}}X) - \mathbb{E}(\mathbf{1}_{\{Y_i \in I_k\}}\mu_{ik})) \otimes \mu_{ik} \\ &= (p_{ik}\mu_{ik} - p_{ik}\mu_{ik}) \otimes \mu_{ik} \\ &= 0\end{aligned}$$

and, similarly, $\mathbb{E}(\mu_{j\ell} \otimes (\mathbf{1}_{\{Y_j \in I_\ell\}}(X - \mu_{j\ell}))) = 0$. Since we obviously have the equality $\mathbb{E}(u_{ik}) \otimes \mathbb{E}(u_{j\ell}) = p_{ik}p_{j\ell} \mu \otimes \mu$, we deduce the required equality: $\sigma_{ikj\ell} = V^{(ikj\ell)} + p_{ik}p_{j\ell}(V - V^{(ik)} - V^{(j\ell)})$. \square

Acknowledgements. The authors thank the Agence Universitaire de la Francophonie (AUF) for financial support, and Professor David Bekolle for academic support.

References

- Allaire, J. and Lepage, Y., 1990. Tests de l'absence de liaison entre plusieurs vecteurs aléatoires pour les distributions elliptiques, *Statist. Anal. données*, **15**, 21-46.
- Anderson, T.W., 1984. An introduction to multivariate statistical analysis, Wiley.
- Bakirov, N.K., Rizzo, M.L. and Székely, G.J., 2006. A multivariate nonparametric test of independence, *J. Multivariate Anal.*, **97**, 1742-1756.
- Billingsley, P., 1968. *Convergence of probability measures*, Wiley.
- Bilodeau, M. and Brenner, D., 1999. *Theory of Multivariate Statistics*, Springer.
- Cléroux, R., Lazraq, A. and Lepage, Y., 1995. Vector correlation based on ranks and a nonparametric test of no association between vectors, *Commun. Statist.-Theory Meth.* **24**, 713-733.
- Dauxois, J., Romain, Y. and Viguier, S., 1994. Tensor products and statistics, *Linear Algebra Appl.*, **210**, 59-88.
- Gieser, P.W. and Randles, R.H., 1997. A Nonparametric Test of Independence Between Two Vectors, *J. Amer. Statist. Assoc.* **92**, 561-567.
- Li, K.C., 1991. Sliced Inverse Regression for Dimension Reduction, *J. Amer. Statist. Assoc.* **86**, 316-327.
- Mathai, A.M. and Provost, S.B., 1992. *Quadratic forms in random variables: Theory and applications*. Dekker.
- Meintanis, S.G. and Iliopoulos, G., 2008. Fourier methods for testing multivariate independence. *Comput. Statist. Data Anal.* **52**, 1884-1895.
- Sinha, B.K. and Wieand, H.S. 1977. Multivariate Nonparametric Tests for Independence, *J. Multivariate Anal.* **7**, 572-583.
- Székely, G.J., Rizzo, M.L. and Bakirov, N.K. 2007. Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35**, 2769-2794.
- Taskinen, S. Kankainen, A. and Oja, H., 2003. Sign test of independence between two random vectors. *Statist. Probab. Lett.* **62**, 9-21.