Factor-based comparison of k populations

Sylvie Viguier-Pla

Laboratoire de Statistique et Probabilités
U.M.R. - C5583
Université Paul Sabatier
118, route de Narbonne
31062 Toulouse Cedex, France
E-mail: viguier@cict.fr

A generalized canonical analysis is considered in order to build a test statistic for the comparison of factor subspaces of k related populations. This test constitutes a generalization of a test of comparison of two factor subspaces (Dauxois, Romain and Viguier [4]). A first stage of the generalization is made in the case of two populations, with the infinite dimension, and without the hypothesis of sphericity. A numerical example is given to compare the test statistics for $k = 2$, and another example illustrates the $k-$sample case.

1. INTRODUCTION

We call Factorial Analysis (F.A.) a multivariate statistical method that can be operated by a diagonalization of a symmetric matrix or, more generaly, by the spectral analysis of a selfadjoint operator. Considering the same measures for several samples, the same F.A. may be performed on each sample. Then one of the problems which arises is the research of similarities and dissimilarities between the populations where from are extracted samples, and in particular the research for common factors.

When the F.A. is a Principal Components Analysis (P.C.A.), the problem of comparison applies to covariance or correlation matrices. In the normality case, several authors have proposed tests of hypotheses on the structure of k covariance matrices (Flury [7], Jensen and Johansen [8], Krzanowski [9]) or two matrices (Schott [13]). Some of these tests, as those of equality and proportionnality, are now classic.

Key words: Canonical Analysis, Factorial Analysis, Factor Subspace, comparison test, asymptotic expansion, perturbation theory
AMS subject classifications: 62E20, 62H15, 62H20, 62H25, 62P

In the same context of P.C.A. on normal variables, Chen and Robinson [2] construct a test for the existence of a subspace common to two factor subspaces. Dauxois, Romain and Viguier ([4]) show that this way of building a test can be applied more generally to any F.A., with only conditions of existence of moments for the involved variables.

A reminder of the previously quoted studies, as well as a software, were proposed (Romain and Viguier [12]) with the aim of supplying a coherent means of joint use of these main tests. All the tests developped later were integrated into this software, which is exposed, with examples, in Viguier-Pla [14].

When $k > 2$, the research of common subspaces is only treated in the P.C.A. case, under normality assumption. Meredith [10, 11] discussed about models for factorial invariance. One of his procedures has been retained by Chen and Robinson [1] who propose a statistic, give its asymptotic distribution and an appropriate test based on it for the research of common factors. Our subject is to propose a test adapted to this research in a more general context, that is for any F.A.. The test statistics are obtained by using an appropriate generalization of the Canonical Analysis (C.A.) of two subspaces.

As the levels of the development are important to understand the test elaboration, the statistic which is of our interest appears after the introduction of several intermediate statistics. This final statistic is a chi-square statistic for the test to be easy to be applied.

The first stage of this generalization consists of a rewriting of the test for $k = 2$, in a more general context, that is without the hypotheses of finite dimension and sphericity, hypotheses which have been made by Chen and Robinson [2] and Dauxois, Romain and Viguier [4].

An example illustrates each of the two test statistics resulting from these generalizations, that have been programmed in addition to the software of Viguier-Pla [14].

2. THE MATHEMATICAL CONTEXT

In the whole following text, the notation are almost identical as that of Dauxois, Romain and Viguier [4]. This paper has been written such that it can be read independently. Nevertheless, supplementary details can be found in this former paper.

2.1. The canonical analysis of k factor subspaces

Let $\mathcal{B}_H$ be the Borel $\sigma-$field of a real separable Hilbert space H, $\sigma_\infty(H)$ (resp. $\sigma_2(H)$) the space of the compact (resp. Hilbert-Schmidt) operators on H, and $\|.\|$ the uniform norm. The norm $\|.\|_2$ is associated with the inner product $<.,.>_2$ defined by:

$$<.,.>_2 : \ (A,B) \in (\sigma_2(H))^2 \mapsto \operatorname{tr} A\, B^* \ ,$$

where B^* is the adjoint operator of B, and tr denotes the trace.

In practice, we study a same random variable (r.v.) X with values in H, relatively to k populations Ω_i, through a same random function. This can be formalized as follows.

Let (X,Z) be a pair of r.v.'s defined on the probability space $(\Omega,\mathcal{A},r)$ taking values in $H\times\mathcal{I}$, where $\mathcal{I}$ is a finite subset $\{1,\ldots,k\}$ of $\mathbb{N}^*$. The sets $\Omega_i=[Z=i]$, $i\in\mathcal{I}$, form a partition of Ω. Let $\mathcal{A}_i$ be the trace of $\mathcal{A}$ on Ω_i and r_i the probability $r/r(\Omega_i)$; then X_i may be considered as a r.v. defined on $(\Omega_i,\mathcal{A}_i,r_i)$, distributed as $X_{|[Z=i]}$, which we suppose known.

In the sequel, the index i, that systematically appears first whenever several indices will be necessary, indicates an element related to the Ω_i.

We assume that the X square norm is $r-$integrable, what implies that the X_i square norm is r_i-integrable. The k F.A.'s processed on the X_i's are obtained through the spectral analysis of the compact operators T_i's, which are functions of the X_i's. For example, if the F.A. is a P.C.A., T_i can be the Gramm operator $\mathbb{E}(X_i\otimes X_i)$.

Let $(\lambda_{ij})_{j\in J_i}$ be the strictly decreasing sequence of eigenvalues of T_i. As T_i is compact, the set of indices J_i is countable. We note P_{ij} the orthogonal projector onto the eigenspace E_{ij} associated with λ_{ij}, of multiplicity order p_{ij}: $\dim E_{ij}=\operatorname{tr}P_{ij}=p_{ij}$.

A factor subspace W_i of T_i is defined as a direct sum $\bigoplus_{j\in J_{W_i}} E_{ij}$, where J_{W_i} is a finite subset of $J_i-\{0\}$; we note p_i its dimension $\sum_{j\in J_{W_i}} p_{ij}$, P_i the eigenprojector $\sum_{j\in J_{W_i}} P_{ij}$ from H onto W_i, and I the identity operator on H.

When $k=2$, the research of common factors to the two F.A.'s is processed by the Canonical Analysis of the Factor Subspaces W_1 and W_2 (named C.A.F.S. by Dauxois, Romain and Viguier [4]). This analysis requires the spectral analysis of $\sigma_i=P_iP_{3-i}P_i$.

When $k\geq 2$, it seems natural to name Generalized Canonical Analysis of Factor Subspaces (G.C.A.F.S.) a canonical analysis of the k subspaces $W_1,\ldots,W_k$. We know (cf. Dauxois and Pousse [3]) that several possibilities exist. Noting that the C.A.F.S. can be obtained by the spectral analysis of $\sigma=P_1+P_2$, we are led to retain what we call now the generalized C.A. (named $k-$C.A. by Dauxois and Pousse [3]) of the W_i's, which we define as follows.

Let $B=P_1+\ldots+P_k-I$. As each of the sets J_{W_i} is finite, B is compact. Then the set of eigenvalues of B^2 is countable. We note it $\Lambda=\{\lambda_l\}_{l\in L}$, and

$\{f_l\}_{l\in L}$ a set of associated orthonormal eigenvectors. The λ_l are supposed decreasingly ordered, except for the null eigenvalue, noted λ_0, if it exists.

Then the (compact) $k-$canonical analysis ($k-$C.A.) is defined by

$$\{\{\lambda_l\},\ \{f_l\}\}_{l\in L}\ .$$

The λ_l are named the canonical coefficients, and f_l the canonical elements. The $k-$C.A. can be obtained by the spectral analysis of other operators, as for example B or $B+I$. This $k-$C.A. is unique (up to the signs of the f_l's) if and only if each λ_l is of multiplicity 1.

2.2. The sampling model

The F.A. of the X_i's are indeed processed on samples of size n_i. Let $n = n_1 + \ldots + n_k$. Each X_i is then associated with a random sample $(X_i^1, \ldots, X_i^{n_i})$, where the n_i r.v.'s are independent and identically distributed (i.i.d.) as X_i, and the F.A. of X_i is estimated by the spectral analysis of the estimator $\widehat{T}_i$ of T_i. The asymptotic properties of this sampling model for the F.A.'s can be found in Dossou-Gbete and Pousse [6].

The estimators relative to samples are noted as the elements which they estimate, with a hat ($\widehat{T}_i$ estimates T_i).

Let r_{X_i} be the image of r_i by X_i, $\mathcal{H}$ be the product space $H^{n_1+\ldots+n_k}$ and $\mathcal{B}_{\mathcal{H}}$ its Borel $\sigma-$field. We consider the sampling structure $(\mathcal{H}, \mathcal{B}_{\mathcal{H}}, \bigotimes_{i=1}^{k} r_{X_i}^{\otimes^{n_i}} = r_{n_1,\ldots,n_k})$ and the mapping

$$\begin{aligned} Y_j \ :\ (\mathcal{H}, \mathcal{B}_{\mathcal{H}}, r_{n_1,\ldots,n_k}) &\ \rightarrow\ (H, \mathcal{B}_H, r_{X_{i(j)}}) \\ (h_1, \ldots, h_{n_1}, h_{n_1+1}, \ldots, h_{n_1+\ldots+n_k}) &\ \mapsto\ h_j, \end{aligned}$$

with $i(j)$ the numero of the i^{th} group where j is in the enumeration $1, \ldots, n_1 + n_2 + \ldots + n_k$. By construction, the Y_j's are independent and their square norm are $r_{n_1,\ldots,n_k}-$integrable.

Proposition 2.1. *$\widehat{T}_i$ is a r.v. from $(\mathcal{H}, \mathcal{B}_{\mathcal{H}})$ into $(\sigma_\infty(H), \|.\|)$, and $(\widehat{T}_1, \ldots, \widehat{T}_k)$ is a family of independent r.v.'s.*

Proof. Set $n_0 = 0$, $\widehat{T}_i$ is a measurable function of $Y^{(i)} = (Y_{n_0+n_1+\ldots+n_{i-1}+1}, \ldots, Y_{n_0+n_1+\ldots+n_i})$ and the r.v.'s $Y^{(1)}, \ldots, Y^{(k)}$ are independent. ∎

In a same way, denoting by $\widehat{W_i}$ the resulting subspace from $\widehat{T}_i$ as W_i is resulting from T_i, and setting $\widehat{P}_i$ the orthogonal projector onto $\widehat{W_i}$, we have the

COROLLARY 2.1. *The $\widehat{W}_i$'s compose a family of independent r.v.'s, and the same occurs for the $\widehat{P}_i$'s.*

2.3. Asymptotic study

In this study, we assume that, for all i,

$$(H_1) \qquad \lim_{n_1,\ldots,n_k\to\infty} \frac{n_i}{n} = \alpha_i \text{ with } 0 < \alpha_i < 1 \text{ and } \alpha_1 + \ldots + \alpha_k = 1 \ .$$

Let $\left((\widehat{\lambda}_{ijh})_{h\in\widehat{H}_{ij}}\right)_{j\in J_{\widehat{W}_i}}$ be the decreasing sequence of eigenvalues of $\widehat{T}_i$ and $\widehat{P}_{ij}$ the orthogonal projector onto the orthogonal sum of eigenspaces associated with the eigenvalues $(\widehat{\lambda}_{ijh})_{h\in\widehat{H}_{ij}}$ converging to λ_{ij}.

Let φ be the application: $B \in \sigma_\infty(H) \mapsto \varphi(B) = B + B^*$, S_{ij} the Moore-Penrose generalized inverse of $T_i - \lambda_{ij} I$: $S_{ij} = (T_i - \lambda_{ij} I)^- = \sum_{l\neq j} \frac{P_{il}}{\lambda_{il}-\lambda_{ij}}$, $\widehat{U}_i = n_i^{1/2}(\widehat{T}_i - T_i)$, and $\widehat{Z}^1_{ij} = S_{ij}\widehat{U}_i P_{ij}$. The asymptotic development of $n_i^{1/2}(\widehat{P}_{ij} - P_{ij})$ ensues from the perturbation theory:

$$n_i^{1/2}(\widehat{P}_{ij} - P_{ij}) = -\varphi(\widehat{Z}^1_{ij}) + n_i^{-1/2}\widehat{A}^1_{ij} + \mathcal{O}(n_i^{-1}) \ . \qquad (2.1)$$

where the expression of $\widehat{A}^1_{ij}$ is not needed explicitely, and $\mathcal{O}$ is the Landau symbol, stochastically defined here, i.e. $\mathcal{O}(\alpha(n)) = \mathcal{O}(\beta(n)) \Leftrightarrow \frac{\alpha(n)}{\beta(n)}$ is almost surely (a.s.) bounded for each n of $\mathbb{N}^*$. As $\frac{n_i}{n}$ is bounded, we also have $\mathcal{O}(\beta(n)) = \mathcal{O}(\beta(n_i))$.

Put $\widehat{Z}_{ij} = (\frac{n}{n_i})^{1/2}\widehat{Z}^1_{ij}$ and $\widehat{A}_{ij} = \frac{n}{n_i}\widehat{A}^1_{ij}$, then (2.1) becomes

$$\widehat{P}_{ij} = P_{ij} - n^{-1/2}\varphi(\widehat{Z}_{ij}) + n^{-1}\widehat{A}_{ij} + \mathcal{O}(n^{-3/2}) \ . \qquad (2.2)$$

Let introduce the following notation:

$$\widehat{P}_i = \sum_{j\in J_{W_i}} \widehat{P}_{ij} \quad , \quad P_i = \sum_{j\in J_{W_i}} P_{ij} \ ,$$

$$\widehat{Q}_i = \sum_{j\in J_i - J_{W_i}} \widehat{P}_{ij} \quad , \quad Q_i = \sum_{j\in J_i - J_{W_i}} P_{ij} \ ,$$

$$\widehat{Z}_i = \sum_{j\in J_i - J_{W_i}} \widehat{Z}_{ij} \quad \text{and} \quad \widehat{A}_i = \sum_{j\in J_i - J_{W_i}} \widehat{A}_{ij} \ .$$

Then, adding the equations (2.2) for $j \in J_i - J_{W_i}$, we obtain

$$\widehat{Q}_i = Q_i - n^{-1/2}\varphi(\widehat{Z}_i) + n^{-1}\widehat{A}_i + \mathcal{O}(n^{-3/2})$$

so

$$\widehat{P}_i = P_i + n^{-1/2}\varphi(\widehat{Z}_i) - n^{-1}\widehat{A}_i + \mathcal{O}(n^{-3/2}) \ . \tag{2.3}$$

We do not need to explicit the operator $\widehat{A}_i$, but only to know the relation that comes from the identification of the two terms in n^{-1} in the equality $(\widehat{P}_i)^2 = \widehat{P}_i$:

$$\widehat{A}_i = P_i\widehat{A}_i - [\varphi(\widehat{Z}_i)]^2 + \widehat{A}_i P_i \ .$$

Moreover, because of the asymptotic normality of

$$(n_i^{1/2}(\widehat{P}_{ij} - P_{ij}))_{j\in J_{W_i}} \ ,$$

the linear function of this vector

$$(n^{1/2}(\widehat{P}_i - P_i))$$

is also asymptotically gaussian centered. It is asymptotically distributed as $\varphi(\widehat{Z}_i)$ and $\varphi(\widehat{Z}_i) + n^{-1/2}\widehat{A}_i + \mathcal{O}(n^{-1})$.

3. THE PARTICULAR CASE OF TWO SUBSPACES ($K = 2$)

3.1. Convergence of the sampling C.A.F.S.

From the fact that the $\widehat{P}_i$'s converge to the P_i's, the sequence of the sampling C.A.F.S.'s (that is here C.A.F.S.'s of $\widehat{W}_1$ and $\widehat{W}_2$) converges to the C.A.F.S. of W_1 and W_2.

Setting $\widehat{\sigma}_1 = \widehat{P}_1\widehat{P}_2\widehat{P}_1$, $\widehat{Z}_{W_i} = \sum_{j\in J_{W_i}} \widehat{Z}_{ij}$ and $\sigma_1 = P_1P_2P_1$, we know that

$$\widehat{\sigma}_1 = \sigma_1 + n^{-1/2}\widehat{G} + \mathcal{O}(n^{-1}) \ ,$$

where

$$\widehat{G} = -P_1P_2\varphi(\widehat{Z}_{W_1}) - P_1\varphi(\widehat{Z}_{W_2})P_1 - \varphi(\widehat{Z}_{W_1})P_2P_1$$

is an asymptotically gaussian centered r.v..

3.2. The test for comparison of two factor subspaces

3.2.1. Construction of the test statistic E_n

Let W_i be the sum of the T_i eigenspaces E_{ij} associated with the r_i largest eigenvalues λ_{ij}. Then $J_{W_i} = \{1, \ldots, r_i\}$ and $J_i = \{1, \ldots, r_i, r_{i+1}, \ldots\}$.

As W_i does not include the eventual eigenspace associated with the null eigenvalue λ_0, its dimension p_i is finite.

Putting

$$\widehat{g} = P_1P_2\varphi(\widehat{Z}_1) + P_1\varphi(\widehat{Z}_2)P_1 + \varphi(\widehat{Z}_1)P_2P_1$$

and

$$\begin{aligned}\widehat{k} &= \varphi(\widehat{Z}_1)\varphi(\widehat{Z}_2)P_1 + \varphi(\widehat{Z}_1)P_2\varphi(\widehat{Z}_1) + P_1\varphi(\widehat{Z}_2)\varphi(\widehat{Z}_1) \\ &\quad -P_1P_2\widehat{A}_1 - P_1\widehat{A}_2P_1 - \widehat{A}_1P_2P_1 \ ,\end{aligned}$$

we have:

$$\widehat{\sigma}_1 = \sigma_1 + n^{-1/2}\widehat{g} + n^{-1}\widehat{k} + \mathcal{O}(n^{-3/2}) \ ,$$

where $\widehat{g} + n^{-1/2}\widehat{k} + \mathcal{O}(n^{-1})$ satisfies to the same properties as $\widehat{G}$ in section 3.1.

Let use the following notation:

- $(\rho_j^2)_{j\in J'}$ (where $J' = \{2, \ldots, s\}$) is the strictly decreasing sequence of the eigenvalues (the value 1 is excluded) of σ_1 (and consequently of $\sigma_2 = P_2P_1P_2$), belonging to $(0, 1)$; then ρ_j is a canonical correlation coefficient of the C.A. of W_1 and W_2. Moreover, we set $\rho_1 = 1$ and $\rho_{s+1} = 0$;
- for all j in J', Q_{ij} is the eigenprojector onto the eigenspace of $\sigma_i = P_iP_{3-i}P_i$ associated with ρ_j^2: $Q_{11} = Q_{21}$ is the projector onto the eventual common subspace of W_1 and W_2, and $Q_{i\,s+1}$ is the projector onto the subspace $\ker\sigma_i$;
- q_j is the multiplicity order of ρ_j^2: $q_j = \operatorname{rank} Q_{ij}$; $q_1 = 0$ and $Q_{i1} = 0$ (resp. $q_{s+1} = 0$ and $Q_{i\,s+1} = 0$) if 1 (resp. 0) is not an eigenvalue of σ_1.

Let us denote $\pi_{1,2}$ (resp. $\pi_{1,2^\perp}$; resp. $\pi_{1^\perp,2}$; resp. $\pi_{1^\perp,2^\perp}$) the orthogonal projector onto $W_1 \cap W_2$ (resp. $W_1 \cap W_2^\perp$; resp. $W_1^\perp \cap W_2$; resp. $W_1^\perp \cap W_2^\perp$). Then

$$\sigma_i = \sum_{j=1}^{s+1} \rho_j^2 Q_{ij} = \sum_{j=1}^{s} \rho_j^2 Q_{ij} \text{ and } P_i = \sum_{j=1}^{s} Q_{ij} + \pi_{i,(3-i)^\perp} .$$

As we previously wrote it, the hypothesis H_0 we wish to test is that of existence of a common subspace to W_1 and W_2. With this notation, it may be written

$$H_0 : \ \dim W_1 \cap W_2 = q_1 \text{ or } \ H_0 : \ \operatorname{tr}\pi_{1,2} = q_1 \ .$$

The test of H_0 against its alternative $\mathrm{H}_1 = \overline{\mathrm{H}}_0$ is based on the asymptotic distribution of a sampling statistic computed from the model described previously. If H_0 is true, then $\widehat{\pi}_{1,2}$ and $\widehat{\pi}_{1,2}\widehat{\sigma}_1$ asymptotically converge, in mean, towards the same limit. The asymptotic distributions of the introduced operators lead to the choice of the statistic

$$E_n = n \operatorname{tr} \widehat{\pi}_{1,2}(I - \widehat{\sigma}_1) \ .$$

3.2.2. Distribution of E_n

The distribution of this statistic has been explicited by Dauxois, Romain and Viguier [4] under the restrictive hypotheses of finite dimension for H and of sphericity (equality) of the eigenvalues λ_{ij}, $j \in J_i - J_{W_i}$. The suppression of these two hypotheses require a new expression of the $\widehat{Z}_i$'s:

$$\widehat{Z}_i = \sum_{j \in J_i - J_{W_i}} \widehat{Z}_{ij} \text{ replaces } \widehat{Z}_i = \widehat{Z}_{i\, r_i+1}, \text{ where } q_{r_i+1} \text{ is finite,}$$

and then a little more complex expression in the development of E_n and of its asymptotic distribution. In a same way, the asymptotic study contains a modified demonstration of the asymptotic normality of the $\widehat{Z}_i$'s, because the P_{ij}'s, $j \in J_i - J_{W_i}$, are not automatically of finite number, nor of finite rank.

The asymptotic development of $\widehat{\pi}_{1,2}$ is obtained in the same way as that of $\widehat{P}_{ij}$:

$$\widehat{\pi}_{1,2} = \pi_{1,2} - n^{-1/2}\varphi(\widehat{z}) + n^{-1}\widehat{B} + \mathcal{O}(n^{-3/2}) \qquad (3.1)$$

with $\widehat{z} = (\sigma_1 - I)^{-}\widehat{g}\pi_{1,2}$ and $\widehat{B} = \pi_{1,2}\widehat{B} + [\varphi(\widehat{z})]^2 + \widehat{B}\pi_{1,2}$.

As $\widehat{Z}_i = (\frac{n}{n_i})^{1/2} \sum_{j \in J_i - J_{W_i}} S_{ij}\widehat{U}_i P_{ij}$ and, for all $j \in J_i - J_{W_i}$, $P_{ij}\pi_{1,2} = 0$, the following proposition of Dauxois, Romain and Viguier [4]:

Proposition 3.1. *We have:*

$$E_n = \|\pi_{1,2}(\widehat{Z}_1 - \widehat{Z}_2)\pi_{1^\perp,2^\perp}\|_2^2 + \mathcal{O}(n^{-1/2}) \ .$$

keeps standing and is demonstrated in a similar manner.

For $i = 1, 2$, $\widehat{Z}_i$ is asymptotically distributed as

$Z_i = \alpha_i^{-1/2} \sum_{j \in J_i - J_{W_i}} S_{ij} U_i P_{ij}$ which is a gaussian centered r.v., and $\widehat{Z}_i$ is of finite rank: $\operatorname{rank} \widehat{Z}_i \leq \operatorname{rank} P_i = p_i$. As $\widehat{Z}_1$ and $\widehat{Z}_2$ are independent by construction, we have the

Proposition 3.2. *Under the H_0 hypothesis, the statistic E_n has the same asymptotic distribution as that of the r.v. $\|\pi_{1,2}(Z_1 - Z_2)\pi_{1^\perp,2^\perp}\|_2^2$, where $\pi_{1,2}(Z_1 - Z_2)\pi_{1^\perp,2^\perp}$ is a centered gaussian r.v. of $\sigma_2(H)$.*

3.2.3. *Computational adaptation of the test statistic*

Under the hypothesis H_0, the statistic E_n is asymptotically distributed as a linear combination of non-necessarily independent chi-square variates. Concretely, this distribution is not easily usable for the test of H_0 (lack of tables for this distribution, uneasy and heavy computation of the test statistic). That is why we prefer to compute a transformation E_n^1 of E_n which, under the hypothesis H_0, is chi-square (χ^2) distributed. In order to explicit the number of degrees of freedom (d.f.) of this distribution, we assume that H is of finite dimension p.

Let V be the covariance operator of $\pi_{1,2}(Z_1 - Z_2)\pi_{1^\perp,2^\perp}$ and $\widehat{V}$ be an estimation of V from the sample elements. Then the statistic

$$E_n^1 = \operatorname{tr} \widehat{\pi}_{1,2}(\widehat{Z}_1 - \widehat{Z}_2)\widehat{\pi}_{1^\perp,2^\perp}\widehat{V}^- \widehat{\pi}_{1^\perp,2^\perp}(\widehat{Z}_1 - \widehat{Z}_2)\widehat{\pi}_{1,2}$$

is asymptotically χ^2 distributed with $\nu = \operatorname{rank} V = \operatorname{rank} \pi_{1,2} \times \operatorname{rank} \pi_{1^\perp,2^\perp} = q_1(p + q_1 - p_1 - p_2)$ d.f..

Generally, the term $\widehat{Z}_i$ cannot be computed, because the quantity $\widehat{U}_i = \sqrt{n}(\widehat{T}_i - T_i)$ cannot be evaluated. But noting that

$$\begin{aligned} \pi_{1,2}\widehat{P}_i\pi_{1^\perp,2^\perp} &= \pi_{1,2}(P_i + n^{-1/2}\widehat{Z}_i)\pi_{1^\perp,2^\perp} + \mathcal{O}(n^{-1}) \\ &= n^{-1/2}\pi_{1,2}\widehat{Z}_i\pi_{1^\perp,2^\perp} + \mathcal{O}(n^{-1}) \\ \text{we have } \pi_{1,2}\widehat{Z}_i\pi_{1^\perp,2^\perp} &= n^{1/2}\pi_{1,2}\widehat{P}_i\pi_{1^\perp,2^\perp} + \mathcal{O}(n^{-1/2}) \end{aligned}$$

and the statistic E_n^1 may be approximated by

$$E'_n = n \operatorname{tr} \widehat{\pi}_{1,2}(\widehat{P}_1 - \widehat{P}_2)\widehat{\pi}_{1^\perp,2^\perp}\widehat{V}^- \widehat{\pi}_{1^\perp,2^\perp}(\widehat{P}_1 - \widehat{P}_2)\widehat{\pi}_{1,2} \ ,$$

which, under the hypothesis H_0, is asymptotically distributed as E_n^1.

As V is the covariance operator of a $p \times p$ matrix, it can be written as a four-indices matrix. For computational ease, E'_n can be expressed with $\widehat{V}$ replaced by a two-indices matrix as an estimation of $\operatorname{cov} \operatorname{vec} \pi_{1,2}(Z_1 - Z_2)\pi_{1^\perp,2^\perp} = \pi_{1^\perp,2^\perp} \overset{K}{\otimes} \pi_{1,2}[\widehat{\operatorname{cov}} \operatorname{vec}(Z_1 - Z_2)]\pi_{1^\perp,2^\perp} \overset{K}{\otimes} \pi_{1,2}$ (see Dauxois, Romain and Viguier-Pla [5]), that we note $\widehat{V}_1$.

Then the statistic E'_n can be re-written

$$E'_n = n\,\mathrm{vec}\left[\widehat{\pi}_{1,2}(\widehat{P}_1 - \widehat{P}_2)\widehat{\pi}_{1^\perp,2^\perp}\right]' \widehat{V}_1^{-}\mathrm{vec}\left[\widehat{\pi}_{1,2}(\widehat{P}_1 - \widehat{P}_2)\widehat{\pi}_{1^\perp,2^\perp}\right] .$$

4. TEST OF COMPARISON FOR $K \geq 2$

4.1. Convergence of the sampling G.C.A.F.S.

Let us recall that we consider k factor subspaces W_i and their orthogonal projectors P_i. The generalized canonical analysis of these subspaces can be processed by the spectral analysis of $\sigma = \sum_{i=1}^k P_i$.

Proposition 4.1. *The sequence of the sample G.C.A.F.S.'s (i.e. the G.C.A. of $\widehat{W_1}, \ldots, \widehat{W_k}$) converges to the G.C.A.F.S. of $W_1, \ldots, W_k$.*

Proof. $\widehat{T}_i \underset{n_i\to\infty}{\overset{\text{a.s.}}{\to}} T_i$ so $\widehat{P}_i \underset{n_i\to\infty}{\overset{\text{a.s.}}{\to}} P_i$ and $\widehat{\sigma} = \widehat{P}_1 + \ldots + \widehat{P}_k \underset{n_1,\ldots,n_k\to\infty}{\overset{\text{a.s.}}{\to}} \sigma = P_1 + \ldots + P_k$. ∎

Let $(\mu_j)_{j\in J}$ be the strictly decreasing sequence of the eigenvalues of σ, $(\xi_j)_{j\in J}$ the sequence of the associated eigenprojectors, and $q_j = \mathrm{rank}\,\xi_j$. Let $(\widehat{\mu}_{jl})_{l\in\widehat{J}_j,\, j\in J}$ be the sequence of the eigenvalues of $\widehat{\sigma}$ and $(\widehat{\xi}_{jl})_{l\in\widehat{J}_j,\, j\in J}$ the sequence of the associated eigenprojectors, such that

$$\text{for all } l \in \widehat{J}_j,\ \widehat{\mu}_{jl} \underset{n\to\infty}{\to} \mu_j \text{ and } \sum_{l\in\widehat{J}_j} \widehat{\xi}_{jl} \underset{n\to\infty}{\to} \xi_j \ .$$

The summation of the equalities (2.3) for $i = 1, \ldots, k$ lead to the equality

$$\widehat{\sigma} = \sigma + n^{-1/2}\varphi(\widehat{Z}) - n^{-1}\widehat{A} + \mathcal{O}(n^{-3/2}) \tag{4.1}$$

where $\widehat{Z} = \sum_{i=1}^k \widehat{Z}_i$ and $\widehat{A} = \sum_{i=1}^k \left[\widehat{A}_i P_i + P_i\widehat{A}_i - [\varphi(\widehat{Z}_i)]^2\right]$. As the P_i's are independent, the asymptotic normality of

$$(n^{1/2}(\widehat{P}_i - P_i))_{i=1,\ldots,k}$$

leads to that of

$$n^{1/2}(\widehat{\sigma} - \sigma)$$

and then of $\varphi(\widehat{Z})$.

4.2. A generalization of the test statistic E_n for $k \geq 2$

4.2.1. Deriving the statistic from the generalized canonical analysis

Let $\pi_{1,k}$ be the eigenprojector of σ associated with the eventual eigenvalue k, and $\widehat{\pi}_{1,k}$ the sample projector which converges to $\pi_{1,k}$. If k is not an eigenvalue of σ, then $\pi_{1,k} = 0$. The way it is constructed, $\pi_{1,k}$ is the projector onto the eventual common subspace of $W_1, \ldots, W_k$, that is $W_1 \cap \ldots \cap W_k$. Note that for $k = 2$, $\pi_{1,k} = \pi_{1,2} = Q_{11}$.

The hypothesis of existence of a common subspace of $W_1, \ldots, W_k$ can then be written:

$$H_0 : \ \dim W_1 \cap \ldots \cap W_k = q_1 \text{ or } \ H_0 : \ \ \operatorname{tr} \pi_{1,k} = q_1 \ .$$

This hypothesis will be tested against the alternative $\overline{H}_0$.

If H_0 is true, then the convergence in $\sigma_\infty(H)$ of $\widehat{\sigma}$ to σ and the compactness of $\pi_{1,k}$ imply the convergence in $\sigma_\infty(H)$ of $\widehat{\pi}_{1,k}$ to $\pi_{1,k}$ when $n \to \infty$.

We note that, when $k = 2$, the spectral decompositions of $\widehat{\sigma}$ and $\widehat{\sigma}_1$ may be written

$$\widehat{\sigma} = \sum_{j=1}^{s} \sum_{l=1}^{q_j} \sum_{\varepsilon \in \{-1,1\}} \widehat{\mu}^{\varepsilon}_{jl} \widehat{Q}^{\varepsilon}_{jl} \text{ and } \widehat{\sigma}_1 = \sum_{j=1}^{s} \sum_{l=1}^{q_j} \widehat{\rho}^2_{jl} \widehat{Q}_{1jl} \ ,$$

$$\text{with } \sum_{l=1}^{q_j} \widehat{Q}_{1jl} \underset{n\to\infty}{\to} Q_{1j} \ , \ \sum_{l=1}^{q_j} \widehat{Q}^{\varepsilon}_{jl} \underset{n\to\infty}{\to} Q^{\varepsilon}_j \ ,$$

$$\text{and } \widehat{\mu}^{\varepsilon}_{jl} = 1 + \varepsilon \, \widehat{\rho}_{jl} \ , \text{ where } \varepsilon = -1 \text{ or } 1 \ .$$

As a matter of fact, if $\widehat{u}_{jl}$ is an eigenvector of $\widehat{\sigma}_1$ associated with $\widehat{\rho}^2_{jl}$, then $\widehat{u}_{jl} + \varepsilon \frac{\widehat{P_2} \widehat{u}_{jl}}{\widehat{\rho}_{jl}}$ is an eigenvector of $\widehat{\sigma}$ associated with $1 + \varepsilon \, \widehat{\rho}_{jl}$; conversely, if $\widehat{v}_{jl}$ is an eigenvector of $\widehat{\sigma}$ associated with $\widehat{\mu}^{\varepsilon}_{jl}$, then $\widehat{P}_1 \widehat{P}_2 \widehat{v}_{jl}$ is an eigenvector of $\widehat{\sigma}_1$ associated with $(\widehat{\mu}^{\varepsilon}_{jl} - 1)^2$. The statistic

$$E_n = n \sum_{l=1}^{q_1} (1 - \widehat{\rho}^2_{1l}) = n \operatorname{tr} \widehat{Q}_{11}(1 - \widehat{\sigma}_1) \text{ where } \widehat{Q}_{11} = \sum_{l=1}^{q_1} \widehat{Q}_{11l} \underset{n\to+\infty}{\longrightarrow} \pi_{1,2}$$

may be written, with the eigenvalues of $\widehat{\sigma}$,

$$E_n = n \sum_{l=1}^{q_1} (1 - (\widehat{\mu}^{\varepsilon}_{1l} - 1)^2) = n \sum_{l=1}^{q_1} \widehat{\mu}^{\varepsilon}_{1l} (2 - \widehat{\mu}^{\varepsilon}_{1l}) \ \text{ where } \varepsilon = -1 \text{ or } 1 \ .$$

Now, for all $j = 1, \ldots, s$, $l = 1, \ldots, q_j$ and $\varepsilon = \overset{+}{-} 1$, $\widehat{Q}^{\varepsilon}_{jl}\widehat{\sigma} = \widehat{\mu}^{\varepsilon}_{jl}\widehat{Q}^{\varepsilon}_{jl}$ implies $\widehat{\mu}^{\varepsilon}_{jl} = \operatorname{tr} \widehat{Q}^{\varepsilon}_{jl}\widehat{\sigma}$ and $(\widehat{\mu}^{\varepsilon}_{jl})^2 = \operatorname{tr} \widehat{Q}^{\varepsilon}_{jl}(\widehat{\sigma})^2$; moreover, as $\pi_{1,2} = Q^1_1 = Q_{11}$, we choose $\varepsilon = 1$, and

$$E_n = n \sum_{l=1}^{q_1} \widehat{\mu}^1_{1l}(2 - \widehat{\mu}^1_{1l})) = n \sum_{l=1}^{q_1} \operatorname{tr} \widehat{Q}^1_{1l}\widehat{\sigma}(2\,I - \widehat{\sigma}) = n \operatorname{tr} \widehat{\pi}_{1,2}\widehat{\sigma}(2\,I - \widehat{\sigma}) \ ,$$

so, for k subspaces $W_1, \ldots, W_k$, we may introduce

$$E_n = n \operatorname{tr} \widehat{\pi}_{1,k}\widehat{\sigma}(k\,I - \widehat{\sigma}) \ . \tag{4.2}$$

4.2.2. Convergence of E_n

Let us recall the notation that will be used in the following: P_{ij} (resp. $\widehat{P}_{ij}$) is the projector onto the subspace W_{ij} (resp. $\widehat{W_{ij}}$), associated with the eigenvalue λ_{ij} of T_i, $i = 1, \ldots, k$, $j \in J_{W_i}$; $\widehat{U}_i = \sqrt{n_i}(\widehat{T}_i - T_i)$; $W_i = \bigoplus_{j \in J_{W_i}} W_{ij}$ and $P_i = \sum_{j \in J_{W_i}} P_{ij}$ is the orthogonal projector onto W_i.
Then we obtain from a perturbation of P_i the development (cf. Eq. (2.3)):

$$\widehat{P}_i = P_i + n^{-1/2}\varphi(\widehat{Z}_i) - n^{-1}\widehat{A}_i + \mathcal{O}(n^{-3/2}) \ ,$$

with $\widehat{Z}_i = \sum_{j \in J_i - J_{W_i}} \widehat{Z}_{ij} = \sum_{j \in J_i - J_{W_i}} (\frac{n}{n_i})^{1/2} S_{ij}\widehat{U}_i P_{ij} = \sum_{j \in J_i - J_{W_i}} (\frac{n}{n_i})^{1/2}(T_i - \lambda_{ij}I)^{-}U_i P_{ij}$ and $\widehat{A}_i = P_i\widehat{A}_i - [\varphi(\widehat{Z}_i)]^2 + \widehat{A}_i P_i$.
Let us recall that $\pi_{1,k}$ is the projector onto $W_1 \cap \ldots \cap W_k$. So $\pi_{1,k}$ can also be seen as the eigenprojector of $\sigma = P_1 + \ldots + P_k$ associated with the eigenvalue k. Furthermore, $\widehat{Z}_i\pi_{1,k} = 0$.

A perturbation applied to $\pi_{1,k}$ leads to, in a similar way as for $\pi_{1,2}$, the development (cf. Eq. (3.1)):

$$\widehat{\pi}_{1,k} = \pi_{1,k} - n^{-1/2}\varphi(\widehat{z}) + n^{-1}\widehat{B} + \mathcal{O}(n^{-3/2}) \tag{4.3}$$

with $\widehat{z} = (\sigma - k\,I)^{-}\varphi(\widehat{Z})\pi_{1,k} = (\sigma - k\,I)^{-}\widehat{Z}^*\pi_{1,k}$ because $\widehat{Z}\pi_{1,k} = 0$.
Moreover, as $(\widehat{\pi}_{1,k})^2 = \widehat{\pi}_{1,k}$, $\widehat{B} = \pi_{1,k}\widehat{B} + [\varphi(\widehat{z})]^2 + \widehat{B}\pi_{1,k}$.

Using the previous developments of $\widehat{\pi}_{1,k}$ and $\widehat{\sigma}$, we obtain the

Proposition 4.2. *Under the hypothesis H_0, we have*

$$E_n = k\,tr\pi_{1,k}\left[\sum_{i,l=1}^{k} \widehat{Z}_i\left(\delta_{il}I + (\sigma - k\,I)^{-}\right)\widehat{Z}^*_l\right] + \mathcal{O}(n^{-1/2}) \ ,$$

where δ_{il} is the Kronecker symbol.

Proof. Replace the terms $\widehat{\pi}_{1,k}$ and $\widehat{\sigma}$ of E_n in (4.2) by their developed expression ((4.1) and (4.3)):

$$\begin{aligned}
E_n &= n\,\mathrm{tr}\,\pi_{1,k}\sigma\,(k\,I-\sigma)\\
&\quad + n^{1/2}\mathrm{tr}\,\Big[\pi_{1,k}\Big(\varphi(\widehat{Z})(k\,I-\sigma)-\sigma\varphi(\widehat{Z})\Big)-\varphi(\widehat{z})\sigma(k\,I-\sigma)\Big]\\
&\quad + \mathrm{tr}\,\Big[\pi_{1,k}\Big(\sigma\,\widehat{A}-[\varphi(\widehat{Z})]^2-\widehat{A}(k\,I-\sigma)\Big)\\
&\quad -\varphi(\widehat{z})\Big(-\sigma\varphi(\widehat{Z})+\varphi(\widehat{Z})(k\,I-\sigma)\Big)+\widehat{B}\sigma\,(k\,I-\sigma)\Big]+\mathcal{O}(n^{-1/2})\ .
\end{aligned}$$

Now, by definition of $\pi_{1,k}$, $\pi_{1,k}\sigma = k\,\pi_{1,k}$, so $\mathrm{tr}\,\pi_{1,k}\sigma\,(k\,I-\sigma)=0$ and $\mathrm{tr}\,\Big[\pi_{1,k}\Big(\varphi(\widehat{Z})(k\,I-\sigma)-\sigma\varphi(\widehat{Z})\Big)-\varphi(\widehat{z})\sigma(k\,I-\sigma)\Big]=0$.

In order to simplify the third term of the development of E_n, note that

$$\mathrm{tr}\,\widehat{B}\sigma\,(k\,I-\sigma)=\mathrm{tr}\,[\varphi(\widehat{z})]^2\sigma\,(k\,I-\sigma)=\mathrm{tr}\,\varphi(\widehat{z})\pi_{1,k}\widehat{Z}(\sigma-k\,I)^-\sigma(k\,I-\sigma)\ ,$$

thus

$$\begin{aligned}
E_n &= \mathrm{tr}\,\Big[\pi_{1,k}(k\,\widehat{A}-\widehat{Z}\widehat{Z}^*)\\
&\quad -\varphi(\widehat{z})\Big(-\sigma\varphi(\widehat{Z})+\varphi(\widehat{Z})(k\,I-\sigma)-\pi_{1,k}\widehat{Z}(\sigma-k\,I)^-\sigma\,(k\,I-\sigma)\Big)\Big]\\
&\quad +\mathcal{O}(n^{-1/2})\\
&= \mathrm{tr}\,\Big[\pi_{1,k}(k\,\widehat{A}-\widehat{Z}\widehat{Z}^*)-(\sigma-k\,I)^-\widehat{Z}^*\pi_{1,k}\Big(-k\,\widehat{Z}+\widehat{Z}(k\,I-\sigma)\\
&\quad -\widehat{Z}(\sigma-k\,I)^-\sigma\,(k\,I-\sigma)\Big)+\pi_{1,k}\widehat{Z}(\sigma-k\,I)^-\sigma\,\widehat{Z}^*\Big]+\mathcal{O}(n^{-1/2})\\
&= \mathrm{tr}\,\pi_{1,k}\Big[k\,\widehat{A}+\widehat{Z}\,\big(-I+k\,(\sigma-k\,I)^- -(k\,I-\sigma)(\sigma-k\,I)^-\\
&\quad +(\sigma-k\,I)^-\sigma\,(k\,I-\sigma)(\sigma-k\,I)^- +(\sigma-k\,I)^-\sigma\big)\,\widehat{Z}^*\Big]+\mathcal{O}(n^{-1/2})
\end{aligned}$$

now $\quad (\sigma-k\,I)(\sigma-k\,I)^- = \pi_{\mathrm{Im}\,(\sigma-k\,I)} = I-\pi_{\mathrm{Ker}\,(\sigma-k\,I)} = I-\pi_{1,k}$, so

$$\begin{aligned}
E_n &= \mathrm{tr}\,\pi_{1,k}\Big[k\,\widehat{A}+\widehat{Z}\,\big(-I+k\,(\sigma-k\,I)^- +(I-\pi_{1,k})\\
&\quad -(\sigma-k\,I)^-\sigma\,(I-\pi_{1,k})+(\sigma-k\,I)^-\sigma\big)\,\widehat{Z}^*\Big]+\mathcal{O}(n^{-1/2})\\
&= \mathrm{tr}\,\pi_{1,k}\Big[k\,\widehat{A}+\widehat{Z}\,\big(k\,(\sigma-k\,I)^- -\pi_{1,k}+k\,(\sigma-k\,I)\pi_{1,k}\big)\,\widehat{Z}^*\Big]\\
&\quad +\mathcal{O}(n^{-1/2})
\end{aligned}$$

$$E_n \quad = \operatorname{tr} \pi_{1,k} \left[\sum_{i,l=1}^{k} \widehat{Z}_i \left(k \left(\delta_{il} I + (\sigma - k\, I)^- \right) \widehat{Z}_l^* \right] + \mathcal{O}(n^{-1/2})$$

and the proposition is proved. ▮

Moreover, for all $i = 1, \ldots, k$, $\widehat{Z}_i \underset{n\to\infty}{\to} Z_i$. This induces the

Proposition 4.3. *Under the hypothesis H_0, we have*

$$E_n \underset{n\to\infty}{\to} E = k\, tr \pi_{1,k} \left[\sum_{i,l=1}^{k} Z_i \left(\delta_{il} I + (\sigma - k\, I)^- \right) Z_l^* \right] .$$

4.3. A simpler test statistic, F_n

4.3.1. Asymptotical equivalence with E_n

The choice of E_n is motivated by the fact that it constitutes a generalization of the case $k = 2$. Nevertheless, a simpler statistic

$$F_n = n \operatorname{tr} \widehat{\pi}_{1,k} (k\, I - \widehat{\sigma}) \ ,$$

which, except a multiplicative coefficient, has the same asymptotic properties as those of E_n, should be more attractive.

Proposition 4.4. *Under the hypothesis H_0, we have*

$$F_n = tr \pi_{1,k} \left[\sum_{i,l=1}^{k} \widehat{Z}_i \left(\delta_{il} I + (\sigma - k\, I)^- \right) \widehat{Z}_l^* \right] + \mathcal{O}(n^{-1/2}) \tag{4.4}$$

and

$$F_n \underset{n\to\infty}{\to} F = tr \pi_{1,k} \left[\sum_{i,l=1}^{k} Z_i \left(\delta_{il} I + (\sigma - k\, I)^- \right) Z_l^* \right] .$$

Proof. The developments of $\widehat{\pi}_{1,k}$ and of $\widehat{\sigma}$ lead to the following development of F_n:

$$F_n \ = n \operatorname{tr} \left[\pi_{1,k} - n^{-1/2} \varphi(\widehat{z}) + n^{-1} \widehat{B} \right] \left[k\, I - \sigma - n^{-1/2} \varphi(\widehat{Z}) + n^{-1} \widehat{A} \right] \\ + \mathcal{O}(n^{-1/2})$$

noting that $\pi_{1,k}(k\,I-\sigma)=0$, $\operatorname{tr}\varphi(\widehat{z})(k\,I-\sigma)=0$ and $\operatorname{tr}\pi_{1,k}\varphi(\widehat{Z})=0$, we have

$$F_n = \operatorname{tr}\left[\widehat{B}(k\,I-\sigma)+\pi_{1,k}\widehat{A}+\varphi(\widehat{z})\varphi(\widehat{Z})\right]+\mathcal{O}(n^{-1/2})\ .$$

Moreover, using the equalities $(\sigma-k\,I)(\sigma-k\,I)^- = I-\pi_{1,k}$ and $\varphi(\widehat{Z})\pi_{1,k} = \widehat{Z}^*\pi_{1,k}$,

$$\begin{aligned}
\bullet\ \operatorname{tr}\widehat{B}(k\,I-\sigma) &= \operatorname{tr}\left[\varphi(\widehat{z})\right]^2(k\,I-\sigma) = \operatorname{tr}\widehat{z}\,\widehat{z}^*(k\,I-\sigma)\\
&= \operatorname{tr}(\sigma-k\,I)^-\varphi(\widehat{Z})\pi_{1,k}\varphi(\widehat{Z})(\sigma-k\,I)^-(k\,I-\sigma)\\
&= \operatorname{tr}(\sigma-k\,I)^-\,\widehat{Z}^*\pi_{1,k}\widehat{Z}(\pi_{1,k}-I)\\
&= -\operatorname{tr}\pi_{1,k}\widehat{Z}(\sigma-k\,I)^-\,\widehat{Z}^*\ ,\\
\bullet\ \operatorname{tr}\pi_{1,k}\widehat{A} &= \sum_{i=1}^{k}\operatorname{tr}\pi_{1,k}\widehat{A}_i = \sum_{i=1}^{k}\operatorname{tr}\pi_{1,k}\widehat{Z}_i\widehat{Z}^*\\
&\text{and}\\
\bullet\ \operatorname{tr}\varphi(\widehat{z})\varphi(\widehat{Z}) &= \operatorname{tr}\left[(\sigma-k\,I)^-\,\widehat{Z}^*\pi_{1,k}+\pi_{1,k}\widehat{Z}(\sigma-k\,I)^-\right]\varphi(\widehat{Z})\\
&= 2\operatorname{tr}\pi_{1,k}\widehat{Z}(\sigma-k\,I)^-\,\widehat{Z}^*\ ,
\end{aligned}$$

what induces

$$\begin{aligned}
F_n &= \operatorname{tr}\left[-\pi_{1,k}\widehat{Z}(\sigma-k\,I)^-\,\widehat{Z}^* + \sum_{i=1}^{k}\operatorname{tr}\pi_{1,k}\widehat{Z}_i\widehat{Z}^* + 2\operatorname{tr}\pi_{1,k}\widehat{Z}(\sigma-k\,I)^-\,\widehat{Z}^*\right]\\
&\quad +\mathcal{O}(n^{-1/2})\\
&= \operatorname{tr}\pi_{1,k}\left[\sum_{i,l=1}^{k}\widehat{Z}_i\left((\sigma-k\,I)^-+\delta_{il}I\right)\widehat{Z}_l^*\right]\pi_{1,k}+\mathcal{O}(n^{-1/2})\ .\\
F_n &\underset{n\to\infty}{\longrightarrow} \operatorname{tr}\pi_{1,k}\left[\sum_{i,l=1}^{k}Z_i\left((\sigma-k\,I)^-+\delta_{il}I\right)Z_l^*\right]\pi_{1,k}\ .
\end{aligned}$$

∎

This proposition leads to the asymptotic equivalence of $k\,F_n$ and E_n:

$$F_n \underset{n\to\infty}{\sim} \frac{1}{k}E_n\ .$$

4.3.2. Quadratic expression of F_n

The operator $k\,I - \widehat{\sigma}$ is clearly positive, so F_n is the trace of a positive operator. In the expression of (4.4), F_n does not appear explicitely positive. If we write it as a quadratic, we shall clarify this point, and furthermore help to compute the test statistic.

Let denote C_n the operator in the expression of F_n such that

$$F_n = \operatorname{tr} C_n + \mathcal{O}(n^{-1/2}) \ ,$$

then, as for $i = 1, \ldots, k$, $\widehat{Z}_i = \widehat{Z}_i P_i^{\perp}$, it comes:

$$C_n = \pi_{1,k} \left[\sum_{i,l=1}^{k} \widehat{Z}_i \left(P_i^{\perp} \delta_{il} P_l^{\perp} - P_i^{\perp} (\sum_{i=1}^{k} P_i^{\perp})^{-} P_l^{\perp} \right) \widehat{Z}_l^* \right] \pi_{1,k} \ .$$

Let $\mathcal{P}^{\perp}$ and $\widehat{Y}$ be respectively the operators such that, for $(x_1, \ldots, x_k)$ in H^k, $\mathcal{P}^{\perp}(x_1, \ldots, x_k) = (P_1^{\perp} x_1, \ldots, P_k^{\perp} x_k)$ and $\widehat{Y}(x_1, \ldots, x_k) = (\widehat{Z}_1 x_1, \ldots, \widehat{Z}_k x_k)$. Moreover, $P^{\perp}$ denotes the operator such that for x in H, $P^{\perp} x = (P_1^{\perp} x, \ldots, P_k^{\perp} x)$.

So C_n can be written

$$C_n = \pi_{1,k} \widehat{Y} \left(\mathcal{P}^{\perp} - P^{\perp *} (P^{\perp} P^{\perp *})^{-} P^{\perp} \right) \widehat{Y}^* \pi_{1,k}$$

Let $A = \mathcal{P}^{\perp} - P^{\perp *}(P^{\perp} P^{\perp *})^{-} P^{\perp}$. This operator is easily seen as the orthogonal projector onto the subspace of $W_1^{\perp} \times \ldots \times W_k^{\perp}$ which is orthogonal to the space generated by the orthogonal projector $P^{\perp *}(P^{\perp} P^{\perp *})^{-} P^{\perp}$.

So $A^2 = A = A^*$, $C_n = \pi_{1,k} \widehat{Y} A (\pi_{1,k} \widehat{Y} A)^*$, and it follows

Proposition 4.5. *Let $Y = (Z_1, \ldots, Z_k)$, $\widehat{\zeta} = \pi_{1,k} \widehat{Y} A$ and $\zeta = \pi_{1,k} Y A$. Then,*

$$F_n = tr \widehat{\zeta}\, \widehat{\zeta}^* + \mathcal{O}(n^{-1/2}) \text{ and } F = tr \zeta\, \zeta^* \ .$$

4.3.3. Limit distribution of F_n

As Y is gaussian centered, and the operators $\pi_{1,k}$ and A are not random, ζ is gaussian centered. Let V be the covariance operator of ζ, then $\zeta\,\zeta^*$ is Wishart $W(1, V)$ distributed. Let D be the diagonal matrix of V. It follows

Proposition 4.6. *Under the hypothesis H_0, the statistic F_n is asymptotically distributed as a Wishart $W(1, D)$ r.v..*

As the test based on a Wishart distribution is relatively difficult to apply (Chen and Robinson [1] use simulation for that purpose), we will use the adequate transformation in order to obtain a statistic that, under the hypothesis H_0, will be chi-square distributed. As $\pi_{1,k} Y\,A\,V^{-1/2}$ is centered reduced and normally distributed, $\pi_{1,k} Y\,A\,V^{-} A Y^{*} \pi_{1,k}$ is chi-square distributed with ν d.f., where $\nu = \operatorname{rank} V = \operatorname{rank} \pi_{1,k} \times \operatorname{rank} A = q \times [(k-1)p + q - \sum_{i=1}^{k} p_i]$.
Note that this d.f. is also the generalization of the d.f. for the previous test ($k = 2$). The statistic

$$F_n^1 = \operatorname{tr} \widehat{\pi}_{1,k} \widehat{Y} A\, \widehat{V}^{-} A\, \widehat{Y}^{*} \widehat{\pi}_{1,k}$$

is then, under the hypothesis H_0, asymptotically chi-square distributed with ν d.f.'s.

We cannot compute $\widehat{Z}_i$ in practice (the reasons are the same that for the case $k = 2$, § 3.2.2). So we use the fact that

$$\pi_{1,k} \widehat{P}_i \pi_{1,k}^{\perp} = \pi_{1,k} (P_i + n^{-1/2} \widehat{Z}_i) \pi_{1,k}^{\perp} + \mathcal{O}(n^{-1}) = n^{-1/2} \pi_{1,k} \widehat{Z}_i \pi_{1,k}^{\perp} + \mathcal{O}(n^{-1})$$

$$\text{and then } \pi_{1,k} \widehat{Z}_i \pi_{1,k}^{\perp} = n^{1/2} \pi_{1,k} \widehat{P}_i \pi_{1,k}^{\perp} + \mathcal{O}(n^{-1/2}) \ ,$$

in order to, noting by $\widehat{P}$ the $k-$uplet $(\widehat{P}_1, \ldots, \widehat{P}_k)$, approximate the statistic F_n^1 by

$$F_n' = n \operatorname{tr} \widehat{\pi}_{1,k} \widehat{P} A\, \widehat{V}^{-} A\, \widehat{P}^{*} \widehat{\pi}_{1,k} \ ,$$

which, under the hypothesis H_0, is asymptotically distributed as F_n^1.

4.3.4. Computation of F_n

We need a matrix expression of the operators for the computation of the test statistic.

The matrix expression of the projectors $\widehat{P}_i$ and $\widehat{\pi}_{1,k}$ is direct. But V, as well as the covariance operator of the matrix Y, is a tensor product of matrices, that can be expressed as a matrix with four indices (see Dauxois, Romain and Viguier-Pla [5]).

For this computation, it is more convenient to replace the estimation of V by that of $\operatorname{cov}(\operatorname{vec} \zeta)$, which is a classical matrix with two indices (rows and columns). We have

$$\operatorname{cov}(\operatorname{vec} \zeta) = \operatorname{cov}[\operatorname{vec}(\pi_{1,k} Y\, A)] = (A \overset{K}{\otimes} \pi_{1,k}) \operatorname{cov}(\operatorname{vec} Y)\, (A \overset{K}{\otimes} \pi_{1,k}) \ .$$

The estimation of $\operatorname{cov}(\operatorname{vec} Y)$ derives from that of $\operatorname{cov}(\operatorname{vec} Z_i)$, for each i, using the relation

$$\operatorname{cov}(\operatorname{vec} Y) = \operatorname{diag} \operatorname{cov}(\operatorname{vec} Z_i).$$

5. EXAMPLES

5.1. The case $k = 2$: comparison of the three statistics

We consider the data from Flury [7]: 200 Swiss banknotes have been divided in two groups, 100 genuine and 100 false. The 4 variables taken into account are width of the bill, measured on the left, width of the bill, measured on the right, width of the lower margin, and width of the upper margin. The values of the three test statistics (χ^2_{CR} for the Dauxois, Romain and Viguier [4]'statistic, χ^2_{CV} for an estimation of E'_n and χ^2_{SV} for an estimation of F'_n) and their respective $p-$values p_{CR}, p_{CV}, p_{SV}, are the following:

p_1	p_2	q	d.f.	χ^2_{CR}	p_{CR}	χ^2_{CV}	p_{CV}	χ^2_{SV}	p_{SV}
3	3	3	3	2.59	0.46	2.59	0.46	2.60	0.46
2	2	2	4	20.16	0.0005	13.66	0.008	13.86	0.008
2	2	1	1	0.042	0.84	0.040	0.84	0.040	0.84
1	1	1	3	9.02	0.029	7.44	0.059	8.71	0.03

We can note that for $q = p_i - 1$, the first two results are identical, because this case is not influenced by the sphericity hypothesis. For each test and an appropriate level (say 1%), the three results lead to the same conclusion : the first factor is not common to both groups, but in the subspace of rank 2, one direction is common.

5.2. The case $k \geq 2$

Let consider the well known data of Fishers' Irises. A set of 150 irises is divided into 3 groups according to their species : 50 setosa, 50 versicolor and 50 virginica. The four variables are sepal length, sepal width, petal length, and petal width. The following table exposes some results for the comparison of the 3 groups simultaneously based on the covariance matrices:

p_1	p_2	p_3	q	d.f.	χ^2_{SV}	p_{SV}
1	1	1	1	6	32.738	1.10^{-5}
2	1	1	1	5	4.053	0.5487
3	3	3	2	2	0.1064	0.9482
3	3	3	3	6	29.017	6.10^{-5}
3	3	2	2	4	22.695	0.0001

The common axes are not of the same order in the decreasing sequences of eigenvalues. One direction is common to the subspaces of respective dimensions 2, 1 and 1. Two directions are common to respective subspaces of dimensions 3, 3 and 3. This results are in accordance with the results of the tests of comparison of the groups two by two.

Note that to accept the hypothesis for (p_1, p_2, p_3, q) implies accepting the hypothesis for all (p'_1, p'_2, p'_3, q) such that $p'_i \geq p_i$. Thus all the combinations do not need to be tested for a complete study.

The software, which is programmed with the GAUSS language, is available on the web at the address

http://www.lsp.ups-tlse.fr/Viguier/prepublications.html

REFERENCES

1. K.H. Chen and J. Robinson (1985) The asymptotic distribution of a goodness of fit statistic for factorial invariance. *J. Multivariate Anal.* **17** 76-83.
2. K.H. Chen and J. Robinson (1989) Comparison of factor spaces of two related populations. *J. Multivariate Anal.* **8** 190-203.
3. J. Dauxois and A. Pousse (1975) Une extension de l'analyse canonique. Quelques applications. *Ann. Inst. Henri Poincaré* **XI** 355-379.
4. J. Dauxois, Y. Romain and S. Viguier (1993) Comparison of two factor subspaces. *J. Multivariate Anal.* **44** 160-178.
5. J. Dauxois, Y. Romain and S. Viguier-Pla (1994) Tensor products and Statistics. *Linear Alg. and its Appl. - Special Issue on Linear Algebra and Statistics* **210** 59-88.
6. S. Dossou-Gbete and A. Pousse (1991) Asymptotic study of eigenelements of a sequence of random selfadjoint operators. *Statistics* **22** 479-491.
7. B. N. Flury (1988) *Common Principal Components and Related Multivariate Models.* Wiley, New York.
8. S.T. Jensen and S. Johansen (1988) Estimation of proportional covariances. *Stat. and Prob. Letters* **6** 83-85.
9. W.J. Krzanowski (1984). Principal component analysis in the presence of group structure. *Appl. Statist.* **33** 164-168.
10. W. Meredith (1964) Notes on factorial invariance. *Psychometrika* **29**(2) 177-185.

11. W. Meredith (1964) Rotation to achieve factorial invariance. *Psychometrika* **29**(2) 187-206.

12. Y. Romain et S. Viguier (1989) COMPCOV : un outil de comparaison de plusieurs matrices de covariance. *Stat. et Anal. des Données* **14** 37-52.

13. J.R. Schott (1988) Common principal components subspaces in two groups. *Biometrika* **75** 229-236.

14. S. Viguier-Pla (1994) Factor-based comparison of several populations using COMPAR software. *Computational Statistics* **9** 135-163.