

Closed Likelihood-Ratio Testing Procedures to Assess Similarity of Covariance Matrices

Francesca Greselin and Antonio Punzo

Abstract Different degrees of similarity can be devised among the k covariance matrices Σ_h , referred to k groups, using their spectral decomposition. In this paper we introduce a closed testing procedure allowing for a choice between eight patterns of covariances. The new methodology allows to disclose a richer information on the data underlying structure than the classical existing methods, and also a more parsimonious parameterization. An application on a real data set exemplify the proposal and shows its performances.

Key words: Eigenvalue Decomposition, Proportional Covariance Matrices, Common Principal Components, Closed Testing Procedures, Likelihood-Ratio Tests.

1 Introduction and motivation

This paper extends the study of similarity between k covariance matrices Σ_h , referred to k groups, under the assumption of multivariate normality. Consider p variables measured on n statistical units arising from $k \geq 2$ different groups. Let $x_1^{(h)}, \dots, x_{n_h}^{(h)}$ denote n_h independent observations, for the h th group, drawn from a normal distribution with mean vector μ_h and covariance matrix Σ_h , $h = 1, \dots, k$. Naturally, $\sum_{h=1}^k n_h = n$.

Suppose to be interested in making inference about $\Sigma_1, \dots, \Sigma_k$, with particular emphasis on their degree of similarity. In this paper, following Celeux and Govaert (1995), we develop an analysis of similarity between covariance matrices considering the decomposition $\Sigma_h = \lambda_h \Gamma_h \Delta_h \Gamma_h'$, ($h = 1, \dots, k$), where Δ_h is

Francesca Greselin

Dipartimento di Metodi Quantitativi per le Scienze Economiche ed Aziendali, à di Milano-Bicocca (Italy), e-mail: francesca.greselin@unimib.it

Antonio Punzo

Dipartimento di Economia e Impresa,
Università di Catania (Italy), e-mail: antonio.punzo@unict.it

the scaled ($|\Delta_h| = 1$) diagonal matrix of the eigenvalues of Σ_h sorted in decreasing order, Γ_h is the $p \times p$ orthogonal matrix whose columns are the normalized eigenvectors of Σ_h ordered according to their eigenvalues and $\lambda_h = |\Sigma_h|^{1/p}$. Each component of the right side of (1) has a different geometric interpretation in terms of the group scatters: Γ_h governs the axes *orientation*, Δ_h controls the *shape*, and λ_h denotes the *volume* of the ellipsoids of equal concentration. By allowing some but not all of these quantities to vary between groups, we obtain parsimonious and easily interpretable models which are appropriate to describe various practical situations. The resulting models give raise to the family $\widetilde{\mathcal{M}} = \{EEE, VEE, EVE, EEV, VVE, VEV, EVV, VVV\}$, where the three letters are respectively referred to volume, shape and orientation, and each of them can be equal (E) or variable (V) among groups.

In order to select the covariance structure in $\widetilde{\mathcal{M}}$, we define a closed multiple testing procedure characterized by local likelihood-ratio (LR) tests. Let $\overline{\mathcal{M}} = \widetilde{\mathcal{M}} \setminus \{VVV\}$ be the closure, under intersection, of $\mathcal{M} = \{VVE, VEV, EVV\}$. For each $M \in \overline{\mathcal{M}}$, let us denote by H_0^M the corresponding null hypothesis. Thus, for example, H_0^{EEV} is the null hypothesis referred to EEV. We set model VVV as the benchmark (diagnostic), because it is the most general (less constrained) in $\widetilde{\mathcal{M}}$, requiring $kp(p+1)/2$ parameters. This position allow us to define seven tests, the most *omnibus* as possible, which lay in a hierarchy. Rejection of H_0^M for all $M \in \overline{\mathcal{M}}$, implicitly leads to the “not rejection” of H_1^{VVV} .

Then, a primary concern for an MTP is the choice of a suitable *error rate* to control. We choose the *familywise error rate* (FWER); it is defined as the probability of committing at least one Type I error, and it is mostly employed when the number of elementary hypotheses is moderate, as in our case. We will employ *adjusted p-values* which are the natural counterpart, in the multiple testing framework, of the classical *p-values* (see, e.g., Bretz, *et al.*, 2009). We construct the MTP as a *closed testing procedure* (CTP) (Marcus *et al.*, 1976) because the latter are among the most powerful MTPs that strongly control the FWER at level α . Further, they are a natural choice for our context, because they address a family of hypotheses that is closed under intersection.

Now, to assess the hypothesis H_0^M for $M \in \overline{\mathcal{M}}$, we employ the likelihood-ratio (LR) statistic

$$LR_M = -2 \ln \frac{L_M}{L_{VVV}} \quad (1)$$

that, under H_0^M , by the general theory of LR-tests is asymptotically distributed (when $\min_{h=1, \dots, k} n_h \rightarrow \infty$) as a χ^2 with v_M degrees of freedom given by the difference in the number of (free) parameters η_{VVV} and η_M between VVV and M .

2 Testing for covariance similarity in the Crabs data

The crab data set of Campbell *et al.* (1974) on the genus *Leptograpsus*, consists in a sample of $n = 100$ blue crabs, there being $n_1 = 50$ males (group 1) and $n_2 = 50$

females (group 2), each specimen having $p = 2$ measurements (in millimeters) on the rear width (RW) and the length along the midline (CL) of the carapace. By Mardia's test, the two group-conditional distributions can be considered bivariate normal, while the LR-test of homoscedasticity rejects the null hypothesis at any reasonable significance level, giving a p -value $p_{EEE} = 6.66 \cdot 10^{-15}$.

Details on the decomposition (1) are shown in Table 1, while Fig. 1 displays the scatterplot of RW versus CL, in both groups.

Table 1 Decomposition of Σ_h^{VVV} , $h = 1, 2$, according to equation (1), for the two groups in the crabs data.

Group	Volume	Shape	Orientation
<i>males</i> ($n_1 = 50$)	$\lambda_1^{VVV} = 3.22274$	$\Delta_1^{VVV} = \begin{bmatrix} 17.54242 & 0 \\ 0 & 0.05700 \end{bmatrix}$	$\Gamma_1^{VVV} = \begin{bmatrix} 0.27256 & -0.96214 \\ 0.96214 & 0.27256 \end{bmatrix}$
<i>females</i> ($n_2 = 50$)	$\lambda_2^{VVV} = 2.24620$	$\Delta_2^{VVV} = \begin{bmatrix} 17.82737 & 0 \\ 0 & 0.05609 \end{bmatrix}$	$\Gamma_2^{VVV} = \begin{bmatrix} 0.37796 & -0.92582 \\ 0.92582 & 0.37796 \end{bmatrix}$

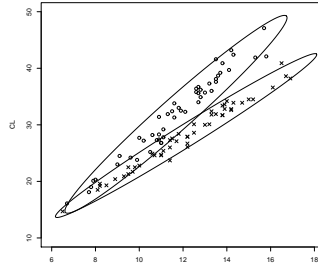


Fig. 1 Scatterplot of variables RW and CL for $n_1 = 50$ males and $n_2 = 50$ females blue crabs (o denotes male and \times female). Ellipses of equal (95%) concentration are also superimposed.

Although the LR-test for model VVV points out heteroscedasticity (with a practically null $p_{EEE} = 6.66 \cdot 10^{-15}$), the scatterplot in Fig. 1 shows strong similarity between volume and shape of the two ellipses; accordingly, in Table 1, sample shape matrices appear to be similar as well as (to a lower extent) sample volumes. The orientation of the two ellipses in Fig. 1 shows a slight difference in the directions of their main axes, attested also in the values along the diagonal of the sample orientation matrices Γ_1^{VVV} and Γ_2^{VVV} in Table 1.

Results in Fig. 2 and Table 2 (see firstly the left six columns) corroborate the aforementioned considerations; in particular, at the 0.05-level, the EEV-model is not rejected since its components (EVV and VEV) are not rejected too. Also, in the comparison between q_{VEV} with q_{EVV} , it is interesting to note the incidence of the former which underlines a stronger degree of similarity between groups in terms of shape. On the contrary, some of the considered likelihood-based ICa (AIC, and also AIC₃, and AICc) lean towards the more complex VEV-model with a loss of one parameter with respect to model EEV.

This paper underlines how some relevant configurations of similarity between covariance matrices Σ_h , referred to different normal groups, can be described by

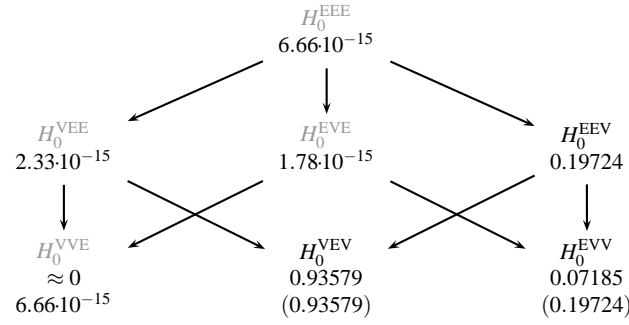


Fig. 2 Unadjusted and adjusted p -values (in round brackets) related to the closed LR-testing procedure applied to the crabs data. Rejected hypotheses, are displayed in gray.

Table 2 Details on the closed LR-testing procedure, and some likelihood-based ICa applied to the crabs data. Bold numbers refer to the “not rejected” model(s) in \mathcal{M} at the 0.05-level (column of the adjusted values q_M), and to the best model for each likelihood-based information criterion.

M	η_M	LR_M	ν_M	p_M	q_M	$-2 \ln L_M$	AIC	BIC	CAIC
EEE	3	69.0889	3	$6.66 \cdot 10^{-15}$		834.6112	840.6112	848.4267	851.4267
VEE	4	67.4247	2	$2.33 \cdot 10^{-15}$		832.9469	840.9470	851.3676	855.3676
EVE	4	67.8934	2	$1.78 \cdot 10^{-15}$		833.4156	841.4156	851.8363	855.8363
EEV	4	3.2466	2	0.19724		768.7689	776.7689	787.1896	791.1896
VVE	5	67.2897	1	≈ 0	$6.66 \cdot 10^{-15}$	832.8119	842.8119	855.8378	860.8378
VEV	5	0.0065	1	0.93579	0.93579	765.5287	775.5287	788.5546	793.5546
EVV	5	3.2403	1	0.07185	0.19724	768.7626	778.7626	791.7884	796.7884
VVV	6					765.5222	777.5223	793.1533	799.1533

considering the three-terms eigenvalue decomposition $\Sigma_h = \lambda_h \Gamma_h \Delta_h \Gamma_h'$. Each of these terms denotes specific geometric characteristics (volume, shape and orientation). This approach leads to eight different models by allowing each of the three terms to be common or not between groups. However, no statistical test to individuate the “correct” model among them exists and, still today, the omnibus Box’s M -test of homoscedasticity (versus heteroscedasticity) is widely used; unfortunately, being omnibus, after a rejection of the null hypothesis, it leaves the practitioner without any more information. In this paper such a gap has been covered by providing a closed testing procedure, using local likelihood-ratio tests, to assess the choice.

References

- Campbell, N. A., and Mahon, R. J. (1974) A graphical approach to sequentially rejective multiple test procedures. *Aust. J. Zool.*, 22 (3), 417–425.
- Celeux, G., and Govaert, G. (1995) Gaussian parsimonious clustering models. *Pattern Recogn.*, 28 (5), 781–793.
- Bretz, F., Maurer, W., Brannath, W., and Posch, M. (2009) A graphical approach to sequentially rejective multiple test procedures. *Stat. Med.*, 28 (4), 586–604.
- Marcus, R., Peritz, E., and Gabriel, K. R. (1976) On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*, 63 (3), 655–660.