Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management
Detroit, Michigan, USA, August 10 - 14, 2020

Inference for Isolated Matrices and Structured Families of Matrices

Cristina Dias
School of Technology and Management, Polytechnic Institute of Portalegre, and Center of Mathematics and Applications (CMA), Portugal
cpsd@ipportalegre.pt

Carla Santos
CMA -Center of Mathematics and its Applications -FCT- New University of Lisbon and Polytechnic Institute of Beja
Campus da Caparica, 2829-516 Caparica, Portugal
carla.santos@ipbeja.pt

João Tiago Mexia
Department of Mathematics and CMA - Center of Mathematics and its Applications -FCT New University of Lisbon
Campus da Caparica, 2829-516 Caparica, Portugal
jtm@unl.pt

Abstract

In this article we show how to use $\boldsymbol{vec}$ type operators to validate models for symmetric stochastic matrices. $\boldsymbol{vec}$ -type operators are operators that they also match vectors to matrices, that is, they allow to reorganize some elements of a matrix into a column vector. The presented formulation allowed us to make inference about the series of study, because the results presented in this work can be applied to the matrices of Hilbert-Schmidt products, that are very important matrices in the first phase of the STATIS methodology. Thus, the operators allowed us to present results that allow inference to be made for models for symmetric stochastic matrices. These models provided the basis for making inference for isolated matrices and structured families of matrices. In particular, we consider the case in which the matrices correspond to the treatments of base models. Thus, using the results presented, using $\boldsymbol{vec}$ type operators, it is possible to adjust a model with degree s. If $\lambda_s = 0$, the model can be simplified. These results, as already mentioned, can be applied to the matrices of Hilbert-Schmidt products (which are very important in the first phase of the STATIS Methodology), and to cross-product matrices which have an important role in inference.

Keywords

Vec type operator's, Cross product matrices, Analysis of Variance, Models, Structured families.

1. Introduction

We use vec type operators in studying models for symmetric stochastic matrices based on their spectral decomposition. Namely we study k-degree models of the form

$$\boldsymbol{M} = \boldsymbol{\mu} + \overline{\boldsymbol{E}} = \sum_{i=1}^{k} \boldsymbol{\lambda_i \alpha_i \alpha_i^t} + \overline{\boldsymbol{E}}$$

© IEOM Society International

for symmetric stochastic matrices $\boldsymbol{M}$, with mean matrix $\boldsymbol{\mu}$, and $\overline{\boldsymbol{E}}$ is a symmetric stochastic matrix with null mean. The models are developed using the spectral analysis of the matrices $\boldsymbol{\mu}$.

The adjustment and validation of the model requires the usage of the vector $\widetilde{\boldsymbol{\beta}}_i = \theta_i \boldsymbol{\gamma}_i$, $i = 1, \dots, k$, which is an estimator of the structure vector $\boldsymbol{\beta}_i = \boldsymbol{\lambda}_i \boldsymbol{\alpha}_i$, $i = 1, \dots, k$ of $\boldsymbol{M}$ with $\lambda_1, \dots, \lambda_k$ the non-null eigenvalues and $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k$ are the eigenvectors, of the matrix $\boldsymbol{M}$.

We intend to introduce a new formulation for the stochastic models of symmetric matrices. In this new formulation, we will use $\boldsymbol{vec}$ type operators to validate models for symmetric stochastic matrices. These operators, who play an important part in the new formulation for these models also facilitate the presentation of these results.

2. Model

The models are developed using the spectral analysis of the matrices $\boldsymbol{\mu}$

$$\boldsymbol{M} = \boldsymbol{\mu} + \overline{\boldsymbol{E}} = \sum_{i=1}^{k} \boldsymbol{\lambda}_i \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i^t + \overline{\boldsymbol{E}} \,,$$

where $\overline{\boldsymbol{E}} = \frac{1}{2}(\boldsymbol{E} + \boldsymbol{E}^t)$ with $\boldsymbol{vec}(\boldsymbol{E}) \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_{n^2})$, as we see in (Dias et al., 2013). The rank of the mean matrix will be the model's degree. The $\boldsymbol{vec}$ type operators are operators are also vectors correspond to matrices. Namely for a symmetric matrix $\boldsymbol{C} = [c_{ij}]$ of order *k*, we have

$$s\_\boldsymbol{vec}(\boldsymbol{C}) = \begin{bmatrix} c_{11} \\ \vdots \\ c_{pp} \\ \vdots \\ c_{kk} \\ \sqrt{2}c_{12} \\ \vdots \\ \sqrt{2}c_{(k-1)k} \end{bmatrix}$$

$$w\_\boldsymbol{vec}(\boldsymbol{C}; p) = \begin{bmatrix} c_{11} \\ \vdots \\ c_{pp} \\ \\ \sqrt{2}c_{12} \\ \vdots \\ \sqrt{2}c_{pk} \end{bmatrix} \quad \text{and} \quad z\_\boldsymbol{vec}(\boldsymbol{C}; p) = \begin{bmatrix} c_{(p+1)(p+1)} \\ \vdots \\ a_{kk} \\ \vdots \\ \sqrt{2}c_{(p+1)(p+2)} \\ \vdots \\ \sqrt{2}c_{(k-1)k} \end{bmatrix},$$

where

- $s_\boldsymbol{vec}(\boldsymbol{C})$is a vector with g components where $g = \frac{k(k+1)}{2}$;
- $w\_\boldsymbol{vec}(\boldsymbol{C}; p)$ is a vector with c_p components where $c_p = g - z_p$;
- $z\_\boldsymbol{vec}(\boldsymbol{C}; p)$ is a vector with z_p components where $z_p = \frac{(k-p)(k-p+1)}{2}$.

We have also

$$\boldsymbol{v_j}(\boldsymbol{C}) = \begin{bmatrix} c_{jj} \\ \\ c_{j(j+1)} \\ \vdots \\ c_{j(k-1)} \\ \vdots \\ c_{jk} \end{bmatrix}$$

© IEOM Society International

which is a column vector whose components are the components of the $j-th$ row vector of matrix $\boldsymbol{C}$ starting at position (j,j). The vector $\boldsymbol{v_j(C)}$ have a_j components where $a_j = k-j+1$, (Harville 1998), (Schott 1997), (Salvador, 2013).

3. Adjustment and validation of the model

Following Oliveira and Mexia (2007) we estimate the structure vector $\boldsymbol{\beta}_i = \boldsymbol{\lambda}_i\boldsymbol{\alpha}_i \ , i = 1, \dots, k$ by the $\widetilde{\boldsymbol{\beta}}_i = \theta_i\boldsymbol{\gamma}_i \ , i = 1, \dots, k$, where $\theta_1, \dots, \theta_n$ are the eigenvalues and the $\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_n$. the eigenvectors of matrix $\boldsymbol{M}$. With $m_1, \dots, m_n$ the column vectors of $\boldsymbol{M}$, we have

$$\widetilde{\boldsymbol{\beta}}_i = \boldsymbol{M}\boldsymbol{\gamma}_i = \boldsymbol{M^t}\boldsymbol{\gamma}_i = \begin{bmatrix} m_1^t\boldsymbol{\gamma}_i \\ \vdots \\ m_n^t\boldsymbol{\gamma}_i \end{bmatrix} = \begin{bmatrix} \boldsymbol{\gamma}_i^t m_1 \\ \vdots \\ \boldsymbol{\gamma}_i^t m_n \end{bmatrix} = (\boldsymbol{I}_n \otimes \boldsymbol{\gamma}_i^t)\begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix}\boldsymbol{Z} \ , i = 1, \dots, k \ ,$$

so $\widetilde{\boldsymbol{\beta}}_i = \theta_i\boldsymbol{\gamma}_i \ , i = 1, \dots, k$ will have mean vector and variance-covariance matrix

$$\begin{cases} E\left(\widetilde{\boldsymbol{\beta}}_i\right) = \left(\boldsymbol{I}_n \otimes \boldsymbol{\gamma}_i^t\right)\boldsymbol{\eta} \ , i = 1, \dots, k \\ \sum\left(\widetilde{\boldsymbol{\beta}}_i\right) = \sigma^2\left(\boldsymbol{I}_n \otimes \boldsymbol{\gamma}_i^t\right)\boldsymbol{L}\left(\boldsymbol{I}_n \otimes \boldsymbol{\gamma}_i \right) \ , i = 1, \dots, k \end{cases},$$

If $\boldsymbol{M}$ has degree k, and the $\widetilde{\boldsymbol{\beta}}_1, \dots, \widetilde{\boldsymbol{\beta}}_k$ $[\boldsymbol{\gamma}_1 \dots \boldsymbol{\gamma}_k]$ are good estimators of $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_k$ $[\boldsymbol{\alpha}_1 \dots \boldsymbol{\alpha}_k]$ then, we can write

$$\overline{\boldsymbol{E}} = \boldsymbol{M} - \sum_{\boldsymbol{i=1}}^{\boldsymbol{k}} \boldsymbol{\beta}_i \, \boldsymbol{\alpha}_i^{\boldsymbol{t}} \approx \boldsymbol{M} - \sum_{\boldsymbol{i=1}}^{\boldsymbol{k}} \widetilde{\boldsymbol{\beta}}_i\boldsymbol{\gamma}_i^t \quad = \boldsymbol{M} - \sum_{\boldsymbol{i=1}}^{\boldsymbol{k}} (\boldsymbol{I}_n \otimes \boldsymbol{\gamma}_i^t)\, \boldsymbol{Z}\boldsymbol{\gamma}_i^t$$

and then

$$\boldsymbol{R} = \boldsymbol{vec}(\overline{\boldsymbol{E}}) \approx \ \boldsymbol{Z} - \boldsymbol{vec}\left(\sum_{\boldsymbol{i=1}}^{\boldsymbol{k}}\left((\boldsymbol{I}_n \otimes \boldsymbol{\gamma}_i^t)\, \boldsymbol{Z}\right)\boldsymbol{\gamma}_i^t \ \right),$$

with

$$\boldsymbol{vec}\left(\sum_{\boldsymbol{i=1}}^{\boldsymbol{k}}\left((\boldsymbol{I}_n \otimes \boldsymbol{\gamma}_i^t)\, \boldsymbol{Z}\right)\boldsymbol{\gamma}_i^t \ \right) = \sum_{\boldsymbol{i=1}}^{\boldsymbol{k}}(\boldsymbol{\gamma}_i \otimes \boldsymbol{I}_n \otimes \boldsymbol{\gamma}_i^t)\, \boldsymbol{Z} \ .$$

Thus

$$\boldsymbol{R} = (\boldsymbol{I}_{n^2} - \boldsymbol{W})\boldsymbol{Z},$$

with $\boldsymbol{W} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{k}}(\boldsymbol{\gamma}_i \otimes \ \boldsymbol{I}_n \otimes \boldsymbol{\gamma}_i^t)$ and $\sum(\boldsymbol{R}) = \sigma^2(\boldsymbol{I}_{n^2} - \boldsymbol{W})\boldsymbol{L}(\boldsymbol{I}_{n^2} - \boldsymbol{W})^t$.
While the vector $\boldsymbol{R}$ can be considered as a residual vector, $\widetilde{\boldsymbol{\beta}}_1, \dots, \widetilde{\boldsymbol{\beta}}_k$ $[\boldsymbol{\gamma}_1 \dots \boldsymbol{\gamma}_k]$ will be the adjusted global structure vector.
Now to check if the model holds it is possible to derive from $\boldsymbol{R}$ a vector that, when the model holds, is normal and homoscedastic, (Areia et al. 2008). We then apply normality tests to that vector.

© IEOM Society International

4. Application to the Inter-structure

In the Joint Table Analysis (ACT), it is customary to have several double-entry data tables, the data that make up these tables are quantitative in nature, which can be obtained in different situations. Each table consists of n rows (objects) and p columns (variables). The various techniques developed for the analysis of this type of tables date back to the 70s and 80s, namely studies of these techniques in Escoufier and L'Hermier des Plantes (1976), Foucart (1981), among others. The various techniques developed allow for a global comparison of all tables, as well as the study of the possible existence of a common structure, which is called inter-structure. The inter-structure is the first stage of the STATIS methodology (STATIS and STATIS Dual method). This methodology was introduced by Hermier des Plantes (1976) and was developed in particular by Lavit (1988). The STATIS methodology uses Euclidean distances between configurations observed in k different situations that we call studies. Whenever, the variables differ throughout the k studies and the number of objects is fixed, there is a proximity between the individuals (objects) we say that we are within the scope of the STATIS method. Whenever the reverse occurs, that is, the number of variables is fixed and objects can change over the course of studies, we say that we are within the scope of the STATIS Dual method. In these two methods, series of studies are studied in which a trio of matrices is considered $(\boldsymbol{X}, \boldsymbol{D}, \dot{\boldsymbol{D}})$, constituted by a data matrix $\boldsymbol{X}$ and two weight matrices, one for objects and one for variables, $\boldsymbol{D}$ and $\dot{\boldsymbol{D}}$ respectively. When the objects are the same for all studies in a series, the same is said of the first type. When the variables are the same for all studies in a series, the same is said for the second type. Escoufier (1978) defined the operators

$$\boldsymbol{A}_i = \boldsymbol{X}_i\dot{\boldsymbol{D}}_i, \boldsymbol{X}_i^t\boldsymbol{D}_i \, , i = 1, \dots, n,$$

and

$$\boldsymbol{B}_i = \boldsymbol{X}_i^t\boldsymbol{D}_i\boldsymbol{X}_i\dot{\boldsymbol{D}}_i \, , i = 1, \dots, n,$$

to represent the studies

$$\left(\boldsymbol{X}_i, \boldsymbol{D}_i, \dot{\boldsymbol{D}}_i\right), i = 1, \dots, n,$$

in the case of series of first and second types, respectively. Let's consider the study of the first type series, putting $\boldsymbol{U} = \boldsymbol{A}$, then we have the matrices $S = [s_{ij}]$,
with

$$s_{ij} = tr\left(\boldsymbol{U}_i\boldsymbol{U}_j^t\right), i = 1, \dots, n,$$

the product of Hilbert-Schmidt operators $\boldsymbol{U}_i \in \boldsymbol{U}_j \, , i = 1, \dots, n$. Being $\theta_1 \geq, \dots, \geq \theta_n$, the eigenvalues and $\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_n$ the corresponding eigenvectors, and according Escoufier (1978), the j-th study is represented by the point whose coordinates are the j-th components of the vectors $\theta_1^{\frac{1}{2}}\boldsymbol{\gamma}_1, \dots, \theta_n^{\frac{1}{2}}\boldsymbol{\gamma}_n$. When these points are along the first axis with the direction of $\boldsymbol{\gamma}_1$, θ_1 it prevails over the other eigenvalues and the series has one-degree structure (Lavit, 1988). In that case it can be admitted having

$$\boldsymbol{S} = \lambda\boldsymbol{\alpha}\boldsymbol{\alpha}^t + \bar{\boldsymbol{E}},$$

therefore, having a symmetric stochastic matrix with degree 1. Thus, it can be admitted that

$$\boldsymbol{U}_i = \boldsymbol{U}_i^0 + \boldsymbol{E}_i \, , i = 1, \dots, n \, ,$$

com $\boldsymbol{U}_i^0 = [u_{ilh}]$ and $\boldsymbol{E}_i = [e_{ilh}]$ with $\boldsymbol{vec}\left(\boldsymbol{E}_i\right) \sim N(0, \, \sigma^2\boldsymbol{I}_{m^2}) \, , i = 1, \dots, n$ (Areia, 2009).

It is also admitted that $\boldsymbol{vec}\left(\boldsymbol{E}_i\right)$ are independent. Thus,

$$S_{ij} = tr\left(\boldsymbol{U}_i\boldsymbol{U}_j^t\right) = \sum_{l=1}^{m}\sum_{h=1}^{m}\left(\boldsymbol{u}_{ilh} + e_{ilh}\right)\left(\boldsymbol{u}_{jlh} + e_{jlh}\right)$$

© IEOM Society International

$$= \sum_{l=1}^{m}\sum_{h=1}^{m}\left(\boldsymbol{u}_{ilh}\boldsymbol{u}_{jlh} + \boldsymbol{u}_{ilh}e_{jlh} + \boldsymbol{u}_{jlh}e_{ilh} + e_{ilh}e_{jlh}\right)$$

$$= S_{ij}^{0} + \sum_{l=1}^{m}\sum_{h=1}^{m}\left(e_{ilh}e_{jlh}\right),\ i,j = 1,\dots,n\ ,$$

with $S_{ij}^{0} = \sum_{l=1}^{m}\sum_{h=1}^{m}\left(\boldsymbol{u}_{ilh}\boldsymbol{u}_{jlh} + \boldsymbol{u}_{ilh}e_{jlh} + \boldsymbol{u}_{jlh}e_{ilh}\right), i,j = 1,\dots,n.$

With $\boldsymbol{S}^{0} = [S_{ij}^{0}]$, $\boldsymbol{vec}(\boldsymbol{S}^{0})$ it will be normally distributed. We can therefore admit that $\boldsymbol{Z} = \boldsymbol{vec}(\boldsymbol{S})$ it will be (approximately) normally distributed as long as we can despised $\sum_{l=1}^{m}\sum_{h=1}^{m}\left(e_{ilh}e_{jlh}\right),\ i,j = 1,\dots,n$. Through simulations carried out see (Dias, 2013, p.85-95), it is shown that when the coefficients of variation $\frac{\sigma_j}{\mu_j}$, of a pair of normal variables $\boldsymbol{X}_j, j = 1,2$ are less than $\frac{1}{17}$, the product $\boldsymbol{X}_1\boldsymbol{X}_2$ can be treated as being normally distributed. So, when the matrix elements $\boldsymbol{U}_1,\dots,\boldsymbol{U}_n$ have coefficients of variation less than $\frac{1}{17}$ we can treat $\mathbf{Z}$ as normal.

With

$$\boldsymbol{K}_{ij} = tr\left(\boldsymbol{U}_i^{0}\boldsymbol{U}_j^{0t}\right) = \sum_{l=1}^{m}\sum_{h=1}^{m}\left(\boldsymbol{u}_{ilh} + \boldsymbol{u}_{jlh}\right), i,j = 1,\dots,n,$$

given the independence and homoscedasticity of $\boldsymbol{vec}\left(\boldsymbol{E}_i\right),\ i = 1,\dots,n$, we have

$$Var\left(S_{ij}^{0}\right) = Var\left(\sum_{l=1}^{m}\sum_{h=1}^{m}\left(\boldsymbol{u}_{ilh} + e_{ilh} + \boldsymbol{u}_{jlh} + e_{jlh}\right)\right) = \sigma^{2}\sum_{l=1}^{m}\sum_{h=1}^{m}\left(\boldsymbol{u}_{ilh}^{2} + \boldsymbol{u}_{jlh}^{2}\right)$$

$$= \sigma^{2}\left(\boldsymbol{K}_{ii} + \boldsymbol{K}_{jj}\right), i \neq j\ ,$$

as well as

$$Var\left(S_{ij}^{0}\right) = Var\left(2\sum_{l=1}^{m}\sum_{h=1}^{m}\left(\boldsymbol{u}_{ilh} + e_{ilh}\right)\right) = 4\sigma^{2}\sum_{l=1}^{m}\sum_{h=1}^{m}\left(\boldsymbol{u}_{ilh}^{2}\right)$$

$$= 4\sigma^{2}\left(\boldsymbol{K}_{ii}\right), i = 1,\dots,n.$$

The remaining elements of the variance-covariance matrix $(\boldsymbol{L})$ are (approximately) given by (Areia, 2009),

$$\begin{cases} Cov\left(S_{ii},S_{ij}\right) = Cov\left(S_{ii},S_{ji}\right) = 2\sigma^{2}\boldsymbol{K}_{ij},\ \ i \neq j. \\ Cov\left(S_{ii},S_{ji}\right) = 0\ \ i \neq j \\ Cov\left(S_{ij},S_{ji}\right) = Var\left(S_{ij}\right)\sigma^{2}\left(\boldsymbol{K}_{ii} + \boldsymbol{K}_{jj}\right), i \neq j \\ Cov\left(S_{ij},S_{il}\right) = Cov\left(S_{ij},S_{li}\right) \approx \sigma^{2}\boldsymbol{K}_{jl},\ \ i \neq j \neq l \\ Cov\left(S_{ij},S_{lu}\right) = 0\ , \qquad i \neq j \neq l \neq u \end{cases}$$

We can then admit (with due approximation) that $\boldsymbol{Z}\sim N(\boldsymbol{\eta},\sigma^{2}\boldsymbol{L})$, with $\boldsymbol{\eta} = \boldsymbol{vec}([\boldsymbol{K}])$ and $\boldsymbol{K} = [k_{ij}]$, since the components of $\boldsymbol{vec}(\boldsymbol{S}^{0})$ being linear functions of a normal vector will have normal joint distribution. Since the S model is a particular case of the base model

$$\boldsymbol{M} = \boldsymbol{\mu} + \overline{\boldsymbol{E}} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{k}}\boldsymbol{\lambda_i}\boldsymbol{\alpha_i}\,\boldsymbol{\alpha_i^{t}} + \overline{\boldsymbol{E}}$$

with $(\lambda_i,\boldsymbol{\alpha_i})$ are the pairs (eigenvalues, eigenvectors) of $\boldsymbol{\mu}$ and $\overline{\boldsymbol{E}}$ is a symmetric stochastic matrix. with (λ_i,α_i) are the pairs (eigenvalues, eigenvectors) of $\boldsymbol{\mu}$ and $\overline{\boldsymbol{E}}$ is symmetric stochastic matrix. To avoid considering a possible large

© IEOM Society International

number of small eigenvalues we assume that $\lambda_1 \geq \cdots \geq \lambda_k$>c. With $\beta_i = \lambda_i \alpha_i, i = 1, \ldots, k$ the information contained in **M** can be condensed in a structured vector

$$\boldsymbol{\beta} = [\boldsymbol{\beta_1^t} \ldots \boldsymbol{\beta_k^t}]^t,$$

and, a sum of square of residues

$$V = \|\boldsymbol{M}\|^2 - \|\boldsymbol{\beta}\|^2,$$

where $\|\quad\|$ indicates Euclidean norm both for matrices and vectors.

In this article we use the $\boldsymbol{vec}$ and other relater operators to carry out inference for structured families of symmetric stochastic matrices $\boldsymbol{M}$. These are obtained through the sum of the respective mean matrix and a symmetric stochastic matrix with null mean. We consider that the **vec** operator of the matrix **E** is normal homoscedastic. The matrices on these families correspond to the treatments of a base design, and the inference is centered on model validation and the action of the factors in the base model on mean matrices.

3. Application to Cross Product Matrices

Given the data matrix $\boldsymbol{X}$ of type $v \times d$, the matrices $\boldsymbol{XX^t}$ and $\boldsymbol{X^tX}$ are the matrices of cross products for row and column vector of $\boldsymbol{X^tX}$. In particular if the column vectors of $\boldsymbol{X}$ are the $\boldsymbol{vec}(\boldsymbol{U}_i), i = 1, \ldots, d$, $\boldsymbol{X^tX}$ will be the matrix of Hilbert-Schmidt products of the matrices $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_d$. Thus, these models directly generalize the previous ones, making it easy to verify that the treatment presented extends directly.
In the case of the matrix $\boldsymbol{XX^t}$ we should have

$$k_{ij} = \sum_{l=1}^{d} x_{il}x_{jl} \ , i, j = 1, \ldots, m \ ,$$

with $\boldsymbol{X} = [x_{ih}]$. For the matrix $\boldsymbol{X^tX}$ we will have

$$k_{ij} = \sum_{l=1}^{m} x_{li}x_{lj} \ , i, j = 1, \ldots, d \ .$$

Note that the matrices $\boldsymbol{XX^t}$ and $\boldsymbol{X^tX}$ have the same non-null eigenvalues $\theta_1 \geq, \ldots, \geq \theta_r > 0$, with $r = Car(\boldsymbol{X}) = Car(\boldsymbol{XX^t}) = Car(\boldsymbol{X^tX})$. Thus, the models for $\boldsymbol{XX^t}$ and $\boldsymbol{X^tX}$ will have the same degree k'. Assuming that $d < v$, the model must be adjusted to $\boldsymbol{X^tX}$ which will be a matrix $d \times d$ rather than $\boldsymbol{XX^t}$ that is $v \times v$. At $\dot{\boldsymbol{\gamma}}_1, \ldots, \dot{\boldsymbol{\gamma}}_{k'}$ first k' eigenvectors for $\boldsymbol{X^tX}$ will correspond, as eigenvectors of $\boldsymbol{XX^t}$, the

$$\boldsymbol{\gamma}_i = \frac{1}{\sqrt{\theta_i}} \boldsymbol{X}\dot{\boldsymbol{\gamma}}_i, i = 1, \ldots, k',$$

that can be used in the inference for the $\boldsymbol{XX^t}$ models and their families.

Since $V = \|\boldsymbol{Y}\|^2$, with V the sum of the squared residuals, and considering $\boldsymbol{Y}$ normal with mean vector and covariance matrix $\sigma^2 \boldsymbol{I}_m$, is independent from the structure vector $\dot{\boldsymbol{\beta}}_i$, with

$$\dot{\boldsymbol{\beta}}_i = (\boldsymbol{X^tX})\boldsymbol{\gamma}_i \ , i = 1, \ldots, k'$$

that is independent of

$$\boldsymbol{\beta}_i = \boldsymbol{X}\dot{\boldsymbol{\beta}}_i \ , i = 1, \ldots, k',$$

so, the results for the model corresponding to $\boldsymbol{X^tX}$ are carried directly to the model corresponding to $\boldsymbol{XX^t}$.

© IEOM Society International

In many cases grade one models can be used for these matrices (Oliveira 2001), (Ramos 2007) and (Areia 2011).

4. Structured Families

These vectorial operators, besides presenting themselves as an important part in the new formulation for these models, also facilitate the presentation of the results. The models for symmetric stochastic matrices are the basis for inference for isolated matrices and for structured families of matrices (Dias 2013), (Dias et al. 2016).
In these families the matrices, all of the same order, correspond to the treatments of base models. Since the matrices have all the same order, we are in the balanced case where we have the same number of degrees of freedom for the error for each treatment.
The ANOVA and related techniques are, in the balanced case, robust techniques for heteroscedasticity and even more for non-normality (Ito 1980), (Scheffé 1959).
The symmetric $n \times n$ matrix $\boldsymbol{W} = [w_{ij}]$ has, besides its $\boldsymbol{vec}$ the $semi\_\boldsymbol{vecS}(\boldsymbol{M})$ with components $w_{1,1}, \dots, w_{n,n}, w_{2,3}, \dots, w_{2,n}, \dots, w_{n-1,n}$.
We use the $semi\_\boldsymbol{vec}$ to validate models for symmetric stochastic matrices. When the predominance of the first eigenvalue is very large, we can adopt a model of degree one of the form

$$\boldsymbol{M} = \boldsymbol{\mu} + \overline{\boldsymbol{E}},$$

where $\boldsymbol{\mu}$ is a mean matrix and $\overline{\boldsymbol{E}}$ is a symmetric stochastic matrix with null mean, such that $\overline{\boldsymbol{E}} = \frac{1}{2}(\boldsymbol{E} + \boldsymbol{E}^t)$ with $\boldsymbol{vec}(\boldsymbol{E}) \sim \boldsymbol{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_{n^2})$ normal homoscedastic.
These models, as we mentioned earlier, have been used successfully in several applications, namely the first phase of the STATIS methodology, the inter-structure. Since Hilbert-Schmidt matrices are matrices of cross products, we can use them for the latter, changing to higher-grade models if degree one models do not fit. In what follows we will base ourselves on presenting the main results that we have for these models.
Simulations (presented next, in the Table 1), show that when the first eigenvalue is sufficiently dominant, we can conclude that the mean matrix as rank one and that the first eigenvalue and eigenvector can be used to estimate the sole non null eigenvalue of the mean matrix and the corresponding eigenvector.
It is intended to generalize this approach, that is, when the first n eigenvalues of the matrix $\boldsymbol{M}$ are sufficiently larger than the others, we can use them to estimate $\lambda_1, \dots, \lambda_r$ and the corresponding eigenvectors to estimate $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_r$.
Thus, with $\theta_1 > \theta_2 > \cdots > \theta_n$ the eigenvalues of the matrix $\boldsymbol{M}$ associated to the eigenvectors $\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_r$, respectively. Thus

$$\boldsymbol{M} = \sum_{i=1}^{\boldsymbol{n}} \theta_i \boldsymbol{\gamma}_i \boldsymbol{\gamma}_i^t.$$

The first r eigenvectors of structure can be estimated by $\widetilde{\boldsymbol{\beta}}_i = \theta_i \boldsymbol{\gamma}_i$, with $i = 1, \dots, n$.
The most interesting cases for these types of models occur when the matrices of this family have the first dominant eigenvalue or the first dominant eigenvalues. In that case, the action of the factors of that design, on the components of the first structure vector or the first structure vectors, depending as it is only the first dominant eigenvalue or the first dominant eigenvalues, respectively (Dias et al. 2012), (Salvador 2013).
Let, $(\theta_i, \boldsymbol{\gamma}_i)\ i = 1, \dots, n$ be the pairs of eigenvalues and eigenvectors for $\boldsymbol{M}$. Then

$$\|\boldsymbol{M}\|^2 = \sum_{i=1}^{n} \theta_i^2 ,$$

when

$$R = \frac{\theta_i^2}{\sum_{l=i+1}^{k} \theta_l^2}$$

If c_i is large we can take the pair $(\theta_i, \boldsymbol{\gamma}_i)\ i = 1, \dots, n$ as estimators of $(\lambda_i, \boldsymbol{\alpha}_i)\ i = 1, \dots, n$, (Oliveira and Mexia 2004).
The values considered for k were: 2, 6, 8, 14, 17 and 40 and the values of R were 12, 5; 50, 100, 200, 600 and 800.
In the construction of the $\boldsymbol{M}$ matrices we took advantage of the symmetry of the problem taking $\boldsymbol{\alpha} = \boldsymbol{\zeta}_1$.

© IEOM Society International

For each pair (k, R), 1000 matrices $\boldsymbol{E}$ were generated with

$$\boldsymbol{vec}(\boldsymbol{E}) \sim N(\boldsymbol{0}, \boldsymbol{I}_{k^2})$$

from which we obtained the matrices

$$\boldsymbol{W} = \lambda \boldsymbol{\zeta}_1 \boldsymbol{\zeta}_1^t + \bar{\boldsymbol{E}},$$

with

$$\lambda = \sqrt{2R},$$

and $\boldsymbol{\alpha} = \boldsymbol{\zeta}_1$. Being $\boldsymbol{\gamma}_1$ the first eigenvector of $\boldsymbol{W}$, since

$$\|\boldsymbol{\zeta}_1\|^2 = \|\boldsymbol{\gamma}_1\|^2 = 1,$$

we will have

$$\begin{aligned} \|\boldsymbol{\zeta}_1 - \boldsymbol{\gamma}_1\|^2 &= (\boldsymbol{\zeta}_1 - \boldsymbol{\gamma}_1)^t (\boldsymbol{\zeta}_1 - \boldsymbol{\gamma}_1) \\ &= \boldsymbol{\zeta}_1^t \boldsymbol{\zeta}_1 - 2\boldsymbol{\zeta}_1^t \boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_1^t \boldsymbol{\gamma}_1 \\ &= 2(1 - \boldsymbol{\gamma}_{11}), \end{aligned}$$

since $\boldsymbol{\gamma}_{11} = \boldsymbol{\zeta}_1^t \boldsymbol{\gamma}_1$, with $\boldsymbol{\gamma}_{11}$ the first component of $\boldsymbol{\gamma}_1$. If $\|\boldsymbol{\gamma}_1\|^2 = 1$, we will have $-1 \leq \boldsymbol{\gamma}_{11} \leq 1$, and

$$\|\boldsymbol{\alpha} - \boldsymbol{\gamma}_1\|^2 = \|\boldsymbol{\zeta}_1 - \boldsymbol{\gamma}_1\|^2,$$

is much smaller as much closer to 1 is $\boldsymbol{\gamma}_{11}$. Thus, we can consider $\boldsymbol{\gamma}_{11}$ as a measure of efficiency of $\boldsymbol{\gamma}$**,** to estimate $\boldsymbol{\alpha}$.

Thus, for each pair (k, R), we present in Table 1 results for batches of 1000 simulations, and the respective mean values and standard deviation for $\|\boldsymbol{\alpha} - \boldsymbol{\gamma}_1\|^2$. From the analysis of the Table 1 we see that, there are many times when the estimated values for c_i fall in the zone where we are led to admit having $\boldsymbol{\gamma} \approx \boldsymbol{\alpha}$ and consequently, $\theta_k \approx \lambda$, and when $R > 200$, we can use θ and $\boldsymbol{\gamma}$ as estimators of λ and $\boldsymbol{\alpha}$ respectively.

Table 1. Mean values and standard deviations for $\|\boldsymbol{\alpha} - \boldsymbol{\gamma}_1\|^2$.

R	$k = 2$		$k = 6$		$k = 8$	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
12.5	0.944	0.181	0.691	0.288	0.501	0.287
50	0.941	0.061	0.908	0.118	0.878	0.128
100	0.986	0.008	0.973	0.010	0.965	0.011
200	0.998	0.002	0.997	0.002	0.996	0.003
600	0,999	0,000	0,999	0,001	0,998	0,001
800	0,999	0,000	0,999	0,000	0,000	0,000
R	$k = 14$		$k = 17$		$k = 40$	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
12.5	0,432	0.296	0,378	0.263	0,417	0.243
50	0,787	0.165	0,891	0.177	0,705	0.223

© IEOM Society International

100	0,943	0.038	0,932	0.051	0,869	0.062
200	0,996	0.012	0,956	0.016	0,889	0.016
600	0,998	0,002	0,997	0,004	0,983	0,002
800	0,999	0,000	0,999	0,000	0,999	0,000

Let us put

$$\bar{\boldsymbol{M}}_h = \boldsymbol{M} - \sum_{i=1}^{h} \theta_i \boldsymbol{\gamma}_i \boldsymbol{\gamma}_i^t,$$

where $m_h = \boldsymbol{S}(\bar{\boldsymbol{M}}_h)$. If $k \leq h$ we will have $\bar{\boldsymbol{M}}_h \approx \boldsymbol{E}$, and so the components of $\boldsymbol{S}(\bar{\boldsymbol{M}}_h)$ can be assumed to be normal with null mean values.
With $Y_1, \dots, Y_g$ $(g = n(n-1))$ the components of $\boldsymbol{S}(\bar{\boldsymbol{M}}_h)$ we are then led to test

$$H_{0k}: Y_1, \dots, Y_g \quad i.i.d. \sim N(0, \sigma^2)\ ,$$

using the statistics (Mexia 1989, 1989b)

$$\mathcal{F} = \frac{(g-1)\left(\sum_{i=1}^{g} Y_i\right)^2}{g\sum_{i=1}^{g} {Y_i}^2 - \left(\sum_{i=1}^{g} Y_i\right)^2}\ ,$$

which when the hypothesis holds, has the central F distribution with 1 and $g-1$ degrees of freedom (Mexia 1995). When H_0 thus not hold, the $Y_1, \dots, Y_g$ will have mean values $\mu_1, \dots, \mu_g$ and the numerator and denominator of F will have the non-centrality parameter

$$\delta_1 = \frac{1}{\sigma^2}\left(\frac{1}{g}\sum_{i=1}^{g} \mu_i\right)^2, \text{ and } \delta_2 = \frac{1}{\sigma^2}\left(\sum_{i=1}^{g} \mu_i^2 - \left(\sum_{i=1}^{g} Y_i\right)^2\right)\ ,$$

hence, there will be alternatives in which δ_1 predominates over δ_2 (δ_2 predominates over δ_1 and in which *F* tends to take larger (smaller) values than when H_0 holds. A numerical study showed that, with $f_{p,1,g-1}$ the quantile for probability p of the central F distribution with 1 and $g-1$ degrees of freedom, the acceptance region.

$$\left[f_{\frac{p}{2},1,g-1}; f_{1-\frac{p}{2},1,g-1}\right],$$

leads to, p level tests with power increasing rapidly with δ_1 and δ_2 (Mexia 1995).
We test this hypothesis for increasing values of k. With $\bar{k}$ the smallest value for which the hypothesis is not rejected we will have the degree of the adjusted model given by

$$\boldsymbol{M}^* = \sum_{i=1}^{\bar{k}} \theta_i \boldsymbol{\gamma}_i \boldsymbol{\gamma}_i^t\ .$$

We can now complete the adjustment testing hypothesis that $\boldsymbol{\gamma}_{\bar{k}+1}, \dots, \boldsymbol{\gamma}_n$, $\bar{k} = 1, \dots, n-1$ are normal homoscedastic with null mean vectors independent between them self's, and from $\boldsymbol{\gamma}_{\bar{k}}$, which will have a mean vector$\boldsymbol{\gamma}_{\bar{k}} \neq 0$, using the test statistic

$$\bar{\mathcal{F}}_{\bar{k}} = (n-\bar{k})\frac{\|\gamma_{\bar{k}}\|^2}{\sum_{h=k+1}^{n}\|\gamma_h\|^2}\ .$$

This F test has n and $(n-\bar{k})n$ degrees of freedom and non-centrality parameter

© IEOM Society International

$$\delta_h = \frac{1}{\sigma^2}\left\|\boldsymbol{\gamma}_{\bar{k}}\right\|^2,$$

and the p level critical value $F_{1-p,n,(n-\bar{k})n}$.

5. Conclusion

In this article we show how to use $\boldsymbol{vec}$ type operators to validate the model. The presented formulation allowed us to make inference about the series of study, because the results presented in this work can be applied to the matrices of the products of Hilbert-Schmidt and cross product matrices. These operators allowed us to present results that allow inference to be made for models for symmetric stochastic matrices. The $\boldsymbol{semi\_vec}$ can be used to determinate the number of non-null eigenvalues of a symmetric matrix $\boldsymbol{M} = \boldsymbol{\mu} + \bar{\boldsymbol{E}}$, and we obtained, using adjusted eigenvalues and eigenvectors, an adjusted model for $\boldsymbol{M}$. These models provided the basis for making inference for isolated matrices and structured families of matrices.

Acknowledgements

This work is funded by National Funds through the FCT - Fundação para a Ciência e a Tecnologia, I.P., under the scope of the project UIDB/00297/2020 (Center for Mathematics and Applications)

References

Areia, A., *Séries emparelhadas de estudo*, Tese de Doutoramento, Universidade de Évora, 2009.

Areia, A., M. M. Oliveira, & J. T. Mexia, *Modelling the Compromise Matrix in STATIS*, *Methodology*, 5, pp. 277-288, 2011.

Dias, C., *Modelos e Famílias de Modelos para Matrizes Estocásticas Simétricas*, Ph.D. Thesis, Évora University, 2013.

Dias, C., Santos S., Varadinov M., & Mexia J.T., ANOVA Like Analysis for Structured Families of Stochastic Matrices, *AIP Proceedings of the 12 International Conference of computational Methods in Sciences and Engineering* (ICCMSE 2016), edited by American Institute of Physics, 1790, pp. 140006-1-140006-3, 2016.

Dias, C., Oliveira M. M, Mexia J. T., Models of symmetric stochastic matrices. *AIP Conf. Proc. 1479, 1666–1669. Numerical Analysis and Applied Mathematics*: International Conference of Numerical Analysis and Applied Mathematics, 19–25, 2012.

Dias, C., Santos C., Varadinov M. and João T. Mexia J. T., ANOVA like analysis for structured families of stochastic matrices. 12th International Conference of Computational Methods in Sciences and Engineering, AIP Conf. Proc. 1790, 140008, 2018.

Escou.er, Y., Le traitement des variables vectorielles, *Biometrics*, 29, pp. 750-760, 1973.

Escoufier, Y. & L'Hermier H., *A Propos de la Comparaison Graphique dês Matrices de Variance*, Biom. J., 20(5), 477-483, 1978.

Foucart, T., *Suites de Tableaux et de Sous-Tableaux*, Revue de Statistique Appliquée, XXIX 2, 31-42, 1981.

Harville, A., *Matrix Algebra from a statistician's perspective*, Springer-Verlag, 1998.

Ito, P. K., *Robustness of Anova and Macanova Test Procedures*, P. R. Krishnaiah (ed), Handbook of Statistics1, Amsterdam: North Holland, pp. 199-236, 1980.

L.Hermier, H., *Structuration des tableaux a trois indices de la statisque*, Thèse de 3ème cycle, Université de Montpellier, 1976.

Lavit, Ch., *Analyse Conjointe de Tableaux quantitatifs*, Collection Méthodes +Programmes, Masson, Paris, 1988.

Lavit, Ch., Escou.er, Y., Sabatier, R., Traissac, P., *The ACT (Statis Method*), Computacional Statistics and Data Analysis. Vol. 18, pp. 97-119, 1994.

Mexia, J., *Simultaneous Confidence Intervals. Generalization of the Scheffé Theore,.*Trabalhos de Investigação, nº2, Departamento de Matemática, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, 1989b.

© IEOM Society International

Mexia, J., *Controlled Heterocedasticity, Quotient Vector Spaces and F Tests for Hypotheses on Mean Vectors*, Trabalhos de Investigação, no.2, Departamento de Matemática, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, 1989.

Mexia, J. T., *Introdução à inferência estatística linear*, Centro de Estudos de Matemática Aplicada -Edição Universitárias Lusófonas, Lisboa, 1995.

Oliveira, M. M., Mexia, J., *F Tests for Hypothesis on the Structure Vectors of Series*, Discussiones Mathematicae.Vol. 19, no. 2, 345-353, 1999.

Oliveira, M. M. and Mexia J. T., AIDS in Portugal: endemic versus epidemic forecasting scenarios for mortality, *International Journal of Forecasting*, 20, pp.131–137, 2004.

Oliveira M.M., Mexia JT, *Modelling series of studies with a common structure*, Computacional Statistics and Data Analysis, N.51, pp 5876—5885, 2007.

Salvador, D., Mexia, J. T., Vec type operators and adjusting models for symmetric stochastic matrices, *7th Workshop on Statistics, Mathematics and Computation and 5th International Conference on Risk Analysis*, Tomar, 2013.

Scheffé, H., The Analysis of Variance, New York: John Wiley & Sons, 1959.

Schott, J. R., *Matrix Analysis for Statistics*, Wiley – Interscience, 1997.

Biographies

Cristina Dias complete a degree in Applied Mathematics (University of Évora, Portugal). Concluded the Master of Applied Mathematics (University of Évora, 2000, Portugal); PHD in Applied Mathematics, (University of Évora, 2013, Portugal); Attended Graduate Diploma in Data Analysis in Social Sciences (University Institute of Lisbon, 2015, Portugal). She began her professional career in 1996 as a teacher of secondary public education and entered university education in the year 2000, where she teaches Linear algebra and analytic geometry; Statistic; Probability and statistics; Quantitative methods; Mathematics; Mathematics Applied to Social Sciences; Operational Research, to Engineering and Business Sciences courses. His current research work is in the areas of Multivariate Data Analysis, Generalized Linear Models, Mixed Linear Models, Computational Statistics, Life Sciences and Environment Statistics, Statistical Modeling in Educational Sciences, Linear Analysis and Linear Inference with five main lines Member of the Portuguese Society of Statistics, Portuguese Society of Mathematics, Member of the Portuguese Operational Research Association.

Carla Santos is an Adjunct Professor at the Department of Mathematics and Physical Sciences of the Polytechnic Institute of Beja and member of Center of Mathematics and Applications, FCT- New University of Lisbon. She is PhD in Mathematics by UBI, MSc. in Applied Mathematics by the University of Évora, Portugal (UE) and BSc. in Mathematics by UE. Her main research interests are in Applied Mathematics and Probability and Statistics, more specifically Statistical Inference in Linear Models. She has published research in these fields and also in Mathematical Education.

João Tiago Praça Nunes Mexia was born in Lisbon in June of 1939 from a Catholic family including his parents and his two sisters. His first degree was in forestry, from Universidade Técnica de Lisboa (Lisbon University of Technology). He concentrated a great part of his activities in 19 Ph.D. supervisions and 12 Ph.D. co-supervisions. His research was centered on Linear Statistical Inference, with 63 papers published in International Journals. As pertaining remarks, we point out the support he received in the first years at FCT/UNL, from the then Head of the Mathematics Department, Professor António César de Freitas. Moreover, his colleagues and students have been a constant encouragement in his work. He has more than 10 co-authors. In the last ten years he has actively cooperated with Polish statisticians, namely Professors Tadeusz Calinski, Stanislaw Mejza and Roman Zmyślony. As a high point in his academic career he was the 'promoter' of the Honoris Causa Degree awarded by Universidade Nova de Lisboa to Professor C. R. Rao in 2006. *Biography of Professor João Tiago Praça Nunes Mexia*. Available from: https://www.researchgate.net/publication/267177055_Biography_of_Professor_Joao_Tiago_Praca_Nunes_Mexia[accessed Jun 13 2020].

© IEOM Society International