DOI: 10.2478/bile-2022-0010

Biometrical Letters
Vol. 59 (2022), No. 2, 141-157

Structures of the covariance matrix: An overview

Malwina Janiszewska

Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Wojska Polskiego 28, PL-60-637 Poznań, Poland,
e-mail: malwina.janiszewska@up.poznan.pl

Summary

In this paper, some multivariate and double multivariate modelling approaches are presented. Moreover, this article provides an overview of the modelling of the structure of the covariance matrix. Furthermore, some methods of covariance structure identification are given.

Key words: covariance matrix, linear structures, non-linear structures, identification methods

1. Introduction

Estimation of the covariance matrix plays a key role in multivariate statistical analysis, which can be applied in many fields such as agronomy, medicine, biology, chemistry, meteorology, physics, economics, genetics, etc. The covariance matrix contains dependence information between characteristics, knowledge of which allows one to draw conclusions and make optimal decisions. The unstructured covariance matrix has relatively many unknown parameters to estimate. Moreover, in many cases, the sample size is too small, which leads to the problem of overparameterization. Therefore, the estimator of an unstructured covariance matrix is singular. In many cases, the data are correlated according to some structure, which we can identify and which allows us to overcome the problem of overparameterization. In addition, the use of an appropriate structure provides an opportunity to analyse data with the use of more precise estimates, which improves statistical inference. Depending on the experiment, various models with specific covariance structure naturally appear. Researchers consider such structures as compound symmetry (cf. Cui et al., 2016), first-order autoregression (cf. Lin et al., 2014), first-order autoregressive moving average (cf. Andrews and Ploberger, 1996), Toeplitz matrices (cf. Niang et al., 2012), banded Teoplitz matrices (cf. Filipiak et al., 2018; John and Mieldzioc, 2020) and symmetric circular Toeplitz matrices (cf. Liang et al., 2020). Separability (Kronecker product) is another structure considered among others by Lu and Zimmerman (2005), Srivastava et al. (2008), Filipiak and Klein (2018), Filipiak et al. (2018), Filipiak et al. (2020). Moreover, authors study different structures of factors in a separable structure such as a block compound symmetry structure; cf. Szatrowski (1976), Szatrowski (1982), Filipiak and Klein (2021). In the papers Janiszewska et al. (2020) and Janiszewska et al. (2022) the authors study block covariance matrices, where off-diagonal blocks with compound symmetry or first-order autoregression structures were considered. In the paper Janiszewska and Markiewicz (2022), banded Toeplitz structures in blocks were used.

In this paper, an experiment where m characteristics are measured on n experimental units is considered. Observations have the form of vectors $\mathbf{x}_i$, $i = 1, \ldots, n$. The observations from different experimental units are independent and have normal distribution

$$\mathbf{x}_i \sim N_m\left(\boldsymbol{\mu}, \boldsymbol{\Omega}\right),$$

where $\boldsymbol{\mu}$ is an unknown mean vector and $\boldsymbol{\Omega}$ is an unknown positive definite (p.d.) covariance matrix. The matrix of all observations $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is normally distributed

$$\mathbf{X} \sim N_{m,n}(\boldsymbol{\mu}\mathbf{1}_n', \boldsymbol{\Omega}, \mathbf{I}_n),$$

where $\mathbf{1}_n$ is an $n-$dimensional vector of ones and $\mathbf{I}_n$ is an identity matrix of order n.

This article focuses on modelling the covariance structure. An unstructured covariance matrix means that there is no pattern at all; for this matrix it is only required that the matrix satisfies the condition of positive definiteness. The advantage of an unstructured covariance matrix is that its estimate is in explicit form. The most common method of estimation is maximum likelihood estimation, where the maximum likelihood estimator (MLE) is the sample covariance matrix $\mathbf{S} = \frac{1}{n}\mathbf{X}\left(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)\mathbf{X}'$. However, observe that the unstructured covariance matrix $\boldsymbol{\Omega}$ has $m(m+1)/2$ unknown parameters to estimate. Having some knowledge about characteristics, we can use an appropriate structure with fewer unknown parameters. The estimation of a covariance matrix with a certain structure may be more or less complex depending on the structure under consideration and the method of estimation.

The selection of the appropriate structure can be based on knowledge of the data or on certain identification methods. The basic approach involves graphical methods, e.g., neural networks or heatmaps (cf. Gilson et al. 2019, Mieldzioc et al. 2021) or the graphical lasso algorithm (cf. Devijver and Gallopin 2018). Algebraic techniques, based on the identification of a structure from a set of potential structures, are widely used in the literature; see for example Cui et al. (2016), Filipiak et al. (2021), Janiszewska et al. (2020), Lin et al. (2014). After choosing the most suitable structure for the considered covariance matrix, the natural procedure is to test hypotheses about covariance structure. The most common tests – the Neyman and Pearson Likelihood Ratio (LR) and Rao's score (RS) – are referred to in the statistical literature (cf. Rao 2005).

The paper is organized as follows. In sections 2 and 3, examples of linear and non-linear structures respectively are presented. Some of the recent developments in the block modelling approach are described in section 4. In section 5, a double multivariate modelling approach is presented, with the proposed structures. In the last section, some identification methods are described.

2. Linear structures

One of the types of commonly used structures is linear structures, defined following Anderson (1973) in Definition 1.

Definition 1. *The covariance matrix* $\mathbf{\Gamma}$ *has a linear structure when it can be expressed as a linear combination of known and symmetric* $m \times m$ *matrices* $\mathbf{G}_0, \ldots, \mathbf{G}_z$ *which are assumed to be linearly independent:*

$$\mathbf{\Gamma} = \sum_{i=0}^{z} \sigma_i \mathbf{G}_i,$$

where $\sigma_i \in \mathbb{R}$ *are unknown parameters. It is also assumed that there exists at least one combination of* $\sigma_0, \ldots, \sigma_z$ *such that* $\mathbf{\Gamma}$ *is a positive definite matrix.*

The linear space $\mathcal{L}$ with the basis $\mathbf{G}_i$, $i = 0, \ldots, z$ is called a linear structure space. A special linear space is a quadratic space $\mathcal{A}$ with the property that $\mathbf{A} \in \mathcal{A}$ implies $\mathbf{A}^2 \in \mathcal{A}$ (cf. Seely, 1971). In addition, a linear space $\mathcal{B}$ with the property that $\mathbf{A}$, $\mathbf{B} \in \mathcal{B}$ implies $\mathbf{AB} = \mathbf{BA} \in \mathcal{B}$ is said to be a quadratic commutative space (cf. Seely, 1971).

Let us assume that $\mathbf{\Gamma}$ is a structured covariance matrix. In the case of a quadratic space structure we can easily get a structured MLE by projecting the matrix $\mathbf{S}$ onto the quadratic space $\mathcal{A}$ (cf. Filipiak et al., 2020); $\widehat{\mathbf{\Gamma}} = \mathbf{P}_{\mathcal{A}}(\mathbf{S}) = \sum_{i=1}^{b} \operatorname{tr}(\mathbf{P}_i\mathbf{S})\mathbf{P}_i$, where $\mathbf{P}_1, \ldots, \mathbf{P}_b$ is an orthonormal basis of the space $\mathcal{A}$. From the fact that $\mathbf{S}$ is p.d. it follows that the resulting estimator is also positive definite – the projection of a p.d. matrix onto a quadratic space maintains its definiteness. In the case when the space structure is not quadratic, restricted maximum likelihood estimation becomes challenging and time-consuming. This is related to the need to use iterative methods and ensure the positive definiteness of the covariance matrix at every step of the algorithm. In an alternative approach, projection of the matrix $\mathbf{S}$ onto a non-quadratic space leads to an explicit estimator. However, if the space structure is not quadratic, this method can lead to estimators that do not satisfy required properties and may even be indefinite; cf. Fuglede and Jensen (2013). In such cases, to restore the positive definiteness of the covariance matrix, the modified shrinkage method can be used (cf. Markiewicz and Mieldzioc, 2022; Janiszewska et al., 2022). The quasi shrinkage method is based on a convex combination of the estimator obtained as a result of the projection $\mathbf{S}$ onto a space structure and an appropriate target matrix from a quadratic subspace of the structure space. The choice of target matrix set affects the properties of the estimators. Thus, to apply this method, characteristics of the quadratic spaces are needed. In the literature some characteristics of commutative quadratic spaces are given (cf. Seely, 1971). Moreover, in the case of a block covariance matrix such characteristics are presented by Janiszewska and Markiewicz (2022).

In the next subsections some quadratic commutative, quadratic and linear space structures are presented.

2.1. Spherical structure

The covariance matrix $\mathbf{\Gamma}$ has a spherical structure if it is proportional to the identity matrix:

$$\mathbf{\Gamma} = \alpha \mathbf{I}_m = \begin{pmatrix} \alpha & 0 & \ldots & 0 \\ 0 & \alpha & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & \ldots & \alpha \end{pmatrix}.$$

In the case of the covariance matrix $\mathbf{\Gamma}$ the components of the observation vector have the same variance and are independent. Note that this is the smallest linear structure; the dimension of the space structure is equal to 1. Thus, there is only one unknown parameter to estimate (i.e. α). A matrix proportional to the identity matrix belongs to all of the structure spaces considered in this paper. In addition, this is a commutative quadratic structure. The spherical structure is often used as the set of the target matrix in the shrinkage method (cf. John and Mieldzioc, 2020).

2.2. Diagonal structure

The covariance matrix $\mathbf{\Gamma}$ has a diagonal structure if it is in the following form:

$$\mathbf{\Gamma} = \operatorname{diag}(\alpha_i) = \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \alpha_m \end{pmatrix}, \quad i = 1, \dots, m.$$

In the case of the covariance matrix $\mathbf{\Gamma}$ the components of the observation vector can have heterogenous variances and are independent. The dimension of the structure space is equal to m. Thus, there are m unknown parameters to estimate (i.e. $\alpha_1, \dots, \alpha_m$). Moreover, this is a commutative quadratic structure. Similarly to spherical structure, diagonal structure of the covariance matrix is used to define the set of the target matrix in the shrinkage method (cf. Hannart and Naveau, 2014; Ikeda et al., 2016).

2.3. Compound symmetry structure

The covariance matrix $\mathbf{\Gamma}$ has the compound symmetry structure (CS) if it is in the following form:

$$\mathbf{\Gamma} = \alpha \mathbf{I}_m + \beta(\mathbf{1}_m \mathbf{1}_m' - \mathbf{I}_m) = \begin{pmatrix} \alpha & \beta & \dots & \beta \\ \beta & \alpha & \ddots & \vdots \\ \vdots & \ddots & \ddots & \beta \\ \beta & \dots & \beta & \alpha \end{pmatrix}.$$

Matrices with this structure have constant variances and constant covariances with each other. This is one of the most important non-trivial quadratic commutative structures. The dimension of the structure space is equal to 2. Thus, the number of unknown parameters for the CS structure is 2 (i.e. α, β). This structure is widely used in the literature. In one early paper (Wilks, 1946), a set of measurements on k equivalent psychological tests was considered, where a covariance matrix with equal diagonal elements and equal off-diagonal elements was applied.

2.4. Banded Toeplitz structure

The covariance matrix $\mathbf{\Gamma}$ with the banded Toeplitz structure has the following form:

$$\mathbf{\Gamma} = \alpha_0 \mathbf{I}_m + \sum_{i=1}^{s} \alpha_i (\mathbf{H}_i + \mathbf{H}_i^{'}) = \begin{pmatrix} \alpha_0 & \alpha_1 & \dots & \alpha_s & 0 & \dots & 0 \\ \alpha_1 & \alpha_0 & \alpha_1 & \dots & \alpha_s & \ddots & \vdots \\ \vdots & \alpha_1 & \alpha_0 & \alpha_1 & & \ddots & 0 \\ \alpha_s & & \ddots & \ddots & \ddots & & \alpha_s \\ 0 & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \alpha_1 \\ 0 & \dots & 0 & \alpha_s & \dots & \alpha_1 & \alpha_0 \end{pmatrix},$$

where $\mathbf{H}_i$ is a binary matrix with 1s on the i-th supradiagonal and all other elements equal to 0. Note that the banded Toeplitz structure is linear and the number of unknown parameters for this structure is $s+1$ (i.e. $\alpha_0, \dots, \alpha_s$). Banded covariance matrices are applied for instance in signal processing, including autoregressive (AR) or moving average (MA) image modelling, covariances of Gauss–Markov random processes (cf. Woods, 1972), or numerical approximations to partial differential equations based on finite difference. Moreover, these matrices are used to model the correlation of cyclostationary processes in periodic time series (cf. Chakraborty, 1998).

2.5. Circular Toeplitz structure

The covariance matrix $\mathbf{\Gamma}$ with the circular Toeplitz (CT) structure has the following form:

$$\mathbf{\Gamma} = \alpha_0 \mathbf{I}_m + \sum_{i=1}^{\lfloor m/2 \rfloor} \alpha_i (\mathbf{H}^i + \mathbf{H}^{i'})$$

if m is odd and

$$\mathbf{\Gamma} = \alpha_0 \mathbf{I}_m + \sum_{i=1}^{m/2-1} \alpha_i (\mathbf{H}^i + \mathbf{H}^{i'}) + \alpha_{m/2} \mathbf{H}^{m/2}$$

if m is even, where $\mathbf{H}$ is a binary matrix with 1s on the first supradiagonal, $h_{p1} = 1$.

Note that the circular structure is quadratic commutative and the dimension of this structure space is equal to $\lfloor m/2 \rfloor$+1, thus the number of unknown parameters is $\lfloor m/2 \rfloor$+1 (i.e. $\alpha_0, \ldots, \alpha_{\lfloor m/2 \rfloor}$). Circular matrices are applied, for example, in cases involving the discrete Fourier transform and in the study of cyclic codes for error correction (cf. Shi et al., 2021).

The banded version of the CT structure is in the following form:

$$\mathbf{\Gamma} = \alpha_0 \mathbf{I}_m + \sum_{i=1}^{k} \alpha_i (\mathbf{H}^i + \mathbf{H}^{i'})$$

with $k < \lfloor m/2 \rfloor$.

Note that the banded circular structure is linear, but not quadratic, and the number of unknown parameters for this structure is k+1 (i.e. $\alpha_0, \ldots, \alpha_k$).

3. Non-linear structures

Beyond linear structures, there are also non-linear structures. In this section some non-linear structures are presented.

3.1. First-order autoregressive structure

The first-order autoregressive structure (AR(1)) is a special case of the autoregressive structure of order r (AR(r)).

The covariance matrix $\mathbf{\Gamma}$ has the autoregressive structure of order one if it is in the following form:

$$\mathbf{\Gamma} = \sigma^2\left(\mathbf{I}_m + \sum_{i=1}^{m-1}\rho^i\mathbf{H}_i\right) = \sigma^2\begin{pmatrix} 1 & \rho & \rho^2 & \dots & & \rho^{m-1} \\ \rho & 1 & \ddots & \ddots & & \rho^{m-2} \\ \rho^2 & \ddots & \ddots & \ddots & & \rho^{m-3} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \rho^{m-2} & & & \ddots & \ddots & \rho \\ \rho^{m-1} & \rho^{m-2} & \rho^{m-3} & \dots & \rho & 1 \end{pmatrix}.$$

Matrices with this structure have homogeneous variances, while the correlations between observations decrease exponentially for successive observations, for example, with distances between time points. Note that this structure is not linear and the number of unknown parameters for the AR(1) structure is 2 (i.e. σ^2, ρ). Matrices with this structure also have a Toeplitz structure. The AR(1) structure is widely used in the literature, e.g. in analysing Gaussian series, and AR(1) models often appear as building blocks in more complex models, for example as a means of including correlated errors in regression (cf. Judge et al. 1982).

3.2. First-order autoregressive moving average structure

The first-order autoregressive moving average structure ARMA(1,1) is a special case of the autoregressive moving average structure of order r, s; (ARMA(r,s)). In some applications, the AR(1) model with the structure of the covariance matrix discussed in the previous subsection and the MA(1) model – the first-order moving average model with the structure of the covariance matrix in the following form

$$\mathbf{\Gamma} = \sigma^2\left(\mathbf{I}_m + \rho_1(\mathbf{H}_1 + \mathbf{H}_1^{'})\right) = \sigma^2\begin{pmatrix} 1 & \rho_1 & 0 & \dots & \dots & 0 \\ \rho_1 & 1 & \rho_1 & \ddots & \dots & 0 \\ 0 & \rho_1 & \ddots & \ddots & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 & \rho_1 \\ 0 & \dots & \dots & 0 & \rho_1 & 1 \end{pmatrix},$$

become cumbersome as means to adequately describe the dynamic structure of the data. Thus, autoregressive moving-average models, which combine the ideas of AR and MA models into a compact form, may be used (cf. Tsay, 2002).

The covariance matrix $\mathbf{\Gamma}$ has a first-order autoregressive moving average structure if it is in the following form:

$$\mathbf{\Gamma} = \sigma^2 \left(\mathbf{I}_m + \gamma \sum_{i=1}^{m-1} \rho^{i-1} \mathbf{H}_i \right) =$$

$$= \sigma^2 \begin{pmatrix} 1 & \gamma & \gamma\rho & \dots & & \gamma\rho^{m-3} & \gamma\rho^{m-2} \\ \gamma & 1 & \ddots & \ddots & & & \gamma\rho^{m-3} \\ \gamma\rho & \ddots & \ddots & \ddots & & & \gamma\rho^{m-4} \\ \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \gamma\rho^{m-3} & & & \ddots & \ddots & & \gamma \\ \gamma\rho^{m-2} & \gamma\rho^{m-3} & \gamma\rho^{m-4} & \dots & & \gamma & 1 \end{pmatrix}.$$

Note that the ARMA(1,1) structure is non-linear and the number of unknown parameters for this structure is 3 (i.e. σ^2, γ, ρ).

The ARMA(1,1) structure is commonly used for covariance structures in time series and multivariate analysis. For example, in the paper Lee et al. (2017), this structure was used for the analysis of lung cancer data.

4. Block multivariate modelling approach

Assume the same experiment as in section 1, but such that the characteristics are divided into two groups, where in the first group there are p characteristics and in the second group there are q characteristics. Thus, the vector of observations $\mathbf{x}_i$, $i = 1, \dots, n$ can be divided into two subvectors, $\mathbf{x}_i = (\mathbf{x}'_{ip}, \mathbf{x}'_{iq})'$. The observations from different experimental units are independent and have normal distribution

$$\mathbf{x}_i \sim N_m(\boldsymbol{\mu}, \mathbf{\Omega}),$$

where $m = p + q$, $\boldsymbol{\mu} = (\boldsymbol{\mu}'_p, \boldsymbol{\mu}'_q)'$ is an unknown mean vector, and $\mathbf{\Omega}$ is an unknown p.d. covariance matrix. The matrix of all observations $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ is normally distributed

$$\mathbf{X} \sim N_{m,n}(\boldsymbol{\mu}\mathbf{1}'_n, \mathbf{\Omega}, \mathbf{I}_n),$$

where $\mathbf{1}_n$ is an $n-$dimensional vector of ones and $\mathbf{I}_n$ is an identity matrix of order n.

Due to the division of the observation vector into two subvectors, the covariance matrix can be presented in the following block form:

$$\boldsymbol{\Omega} = \begin{pmatrix} \boldsymbol{\Omega}_{11} & \boldsymbol{\Omega}_{12} \\ \boldsymbol{\Omega}_{12}' & \boldsymbol{\Omega}_{22} \end{pmatrix}.$$

The dependence between two groups of characteristics is presented in the block $\boldsymbol{\Omega}_{12}$ (cross-covariance matrix). If the hypothesis of the independence of two groups is rejected, then it becomes interesting to study the structure of this dependence.

In the literature, some subblocks of the structures given in the previous sections have been considered. In the paper Janiszewska et al. (2020), the authors study the following structure:

$$\boldsymbol{\Gamma} = \begin{pmatrix} \boldsymbol{\Gamma}_1 & \boldsymbol{\Psi} \\ \boldsymbol{\Psi}' & \boldsymbol{\Gamma}_2 \end{pmatrix}$$

with unstructured diagonal blocks of order p and q respectively, and structured off-diagonal blocks such that $\boldsymbol{\Gamma} \in \mathbb{R}_m^>$. The following forms of matrix $\boldsymbol{\Psi}$ are considered:

- the matrix with all elements equal to δ

$$\boldsymbol{\Psi} = \delta \mathbf{1}_p \mathbf{1}_q', \tag{1}$$

- the matrix of the form

$$\boldsymbol{\Psi} = \delta \mathbf{A} = \delta \begin{pmatrix} \varrho^p & \varrho^{p+1} & \cdots & \varrho^{p+q-1} \\ \varrho^{p-1} & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \varrho^2 & \cdots & \cdots & \cdots \\ \varrho & \varrho^2 & \cdots & \varrho^q \end{pmatrix}. \tag{2}$$

Observe that such forms of off-diagonal blocks correspond with off-diagonal blocks from compound symmetry or first-order autoregression structures. In the paper Janiszewska et al. (2022), estimation of the block matrix $\boldsymbol{\Gamma}$ with off-diagonal blocks given in (1) is presented.

We can also study the dependency structure within groups. In the paper Janiszewska and Markiewicz (2022), the authors study the case of block matrices with separately linearly structured blocks, where they use different banded Toeplitz structures. Moreover, studying the dependency between more than two groups is necessary in many areas, and this will be a direction of further research.

5. Double multivariate modelling approach

Assume an experiment in which q characteristics at t time points (locations) are observed for each of n experimental units. The data collected in this way are called doubly multivariate. Let $\mathbf{X}_i$ for $i = 1, \dots, n$ be independent and identically distributed $(q \times t)$-dimensional observation matrices. Moreover,

$$\mathrm{vec}(\mathbf{X}_i) \sim N_{tq}(\mathrm{vec}\,\mathbf{M}, \mathbf{\Omega}),$$

where $\mathrm{vec}\,\mathbf{M} \in \mathbb{R}^{tq}$, $\mathrm{vec}(\cdot)$ is the operator stacking the columns of a $q \times t$ matrix into a tq-dimensional vector, and $\mathbf{\Omega}$ is an unstructured positive definite matrix of order tq. The number of unknown parameters to be estimated in $\mathbf{\Omega}$ is $tq(tq+1)/2$.

The most natural assumption for doubly multivariate models seems to be the assumption that characteristics are correlated regardless of time points and time points are correlated independently of characteristics. Thus, the covariance matrix is expressed as the following Kronecker product:

$$\mathbf{\Gamma} = \mathbf{\Psi} \otimes \mathbf{\Sigma} \in \mathbb{R}^{>}_{tq \times tq},$$

where $\mathbf{\Psi} \in \mathbb{R}^{>}_{t \times t}$ is the matrix describing the unknown covariance structure between the columns of $\mathbf{X}_i$ (time points), and $\mathbf{\Sigma} \in \mathbb{R}^{>}_{q \times q}$ is the matrix describing the unknown covariance structure between the rows of $\mathbf{X}_i$ (characteristics). This structure is also called a separable structure. Note that this structure is non-linear, and the number of unknown parameters to estimate is $\frac{t(t+1)}{2} + \frac{q(q+1)}{2}$.

In the classical approach, $\mathbf{\Psi}$ and $\mathbf{\Sigma}$ remain unstructured; however, these matrices can have some structure. We can give a generalization of the linear structure given in Definition 1 for a multivariate model (cf. Krishnaiah and Lee, 1974). The matrix $\mathbf{\Gamma}$ has a linear structure when it can be expressed as a linear combination of known $t \times t$ matrices $\mathbf{G}_1, \dots \mathbf{G}_z$ in the following form:

$$\mathbf{\Gamma} = \sum_{i=1}^{z} \mathbf{G}_i \otimes \mathbf{\Sigma}_i,$$

where $\mathbf{\Sigma}_i$ is an unknown $q \times q$ matrix (cf. Timm, 2002). Observe that, if the matrix $\mathbf{\Gamma}$ has the following form:

$$\mathbf{\Gamma} = \sum_{i=1}^{z} \mathbf{M}_i \otimes \mathbf{\Sigma}_i,$$

where $\mathbf{M}_i$ are symmetric idempotent and pairwise orthogonal matrices with $\sum_{i=1}^{z} \mathbf{M}_i = \mathbf{I}_t$ and $\mathbf{\Sigma}_i$ are symmetric matrices, then $\mathbf{\Gamma}$ has a quadratic structure (cf. Filipiak et al. 2020).

For instance, if $\mathbf{M}_1 = \mathbf{I}_t$, then the matrix $\mathbf{\Gamma} = \mathbf{M}_1 \otimes \mathbf{\Sigma}_1$ has a multivariate sphericity structure. In the other case, if $\mathbf{M}_1 = \frac{1}{t}\mathbf{1}_t\mathbf{1}_t'$ and $\mathbf{M}_2 = \frac{1}{\sqrt{t-1}}(\mathbf{I}_t - \frac{1}{t}\mathbf{1}_t\mathbf{1}_t')$, then the matrix $\mathbf{\Gamma} = \mathbf{M}_1 \otimes \mathbf{\Sigma}_1 + \mathbf{M}_2 \otimes \mathbf{\Sigma}_2$ has a multivariate compound symmetry structure.

In the paper Filipiak et al. (2020), the authors use an extension of the classical CS covariance matrix of multivariate data for the double-multivariate case. Moreover, the block exchangeable structure (BCS), being a generalization of the compound symmetry structure into blocks, is considered in the literature; cf. Roy et al. (2016), Filipiak and Klein (2021).

6. Structure identification methods

Graphical diagnostics can be very helpful in selecting and evaluating models for covariance structures. To indicate the most suitable structured covariance matrix for a dataset, visualization of the data with heatmaps can be used. In the paper Mieldzioc et al. (2021), the authors use hierarchical clustering to find groups of similar traits. This method uses a measure of dissimilarity for some number of objects being clustered. Distances between elements are computed using a chosen distance function, and the distances between clusters are obtained using the chosen linkage criterion. Based on visualizations of the correlation matrix after hierarchical clustering, some structures of the covariance matrix can be recognized.

To identify which of the possible structures is the most adequate, identification methods can be used. Suppose $\mathbf{A}$ is a given $m \times m$ p.d. covariance matrix. Let $\mathcal{S}$ be the set of all $m \times m$ p.d. covariance matrices with structure s. Moreover, assume there is a given set of k candidate covariance structures $\{s_1, \ldots, s_k\}$, and let S_i be the set of all covariance matrices with structure s_i.

Thus, the discrepancy between the given covariance matrix $\mathbf{A}$ and the set $M = \cup_{i=1}^{k} S_i$ is defined by:

$$D(\mathbf{A}, M) = \min_{\mathbf{B} \in M} L(\mathbf{A}, \mathbf{B}),$$

where $L(\mathbf{A}, \mathbf{B})$ is a measure of the discrepancy between the two $m \times m$ matrices $\mathbf{A}$ and $\mathbf{B}$ (cf. Cui et al., 2016). Identification of the structure is based on determining the matrix from the set M which is the closest to the matrix $\mathbf{A}$ in the sense of some discrepancy function. In the literature, the following discrepancy functions are considered:

- the entropy loss function

$$f_E(\mathbf{A}, \mathbf{B}) = \mathrm{tr}\left(\mathbf{A}^{-1}\mathbf{B}\right) - \ln\left|\mathbf{A}^{-1}\mathbf{B}\right| - m,$$

- the quadratic loss function

$$f_Q(\mathbf{A}, \mathbf{B}) = \mathrm{tr}\left[\left(\mathbf{A}^{-1}\mathbf{B} - \mathbf{I}_m\right)^2\right],$$

- the Frobenius norm

$$f_F(\mathbf{A}, \mathbf{B}) = ||\mathbf{A} - \mathbf{B}||_F^2 = \mathrm{tr}\left[(\mathbf{A} - \mathbf{B})(\mathbf{A} - \mathbf{B})'\right].$$

The entropy loss function was used in the paper Janiszewska et al. (2020) to identify covariance structure, where consideration was given to a block covariance structure with off-diagonal blocks corresponding to the off-diagonal blocks of CS and AR(1) structures. The entropy loss function is also known as Kullback–Leibler divergence between two probability distributions; cf. Pan and Fang (2002). The structure identifications using the entropy loss function and the quadratic loss function were compared by Mokrzycka (2021). Observe that the entropy loss function and quadratic loss function require nonsingularity of the matrix, thus they sometimes cannot be applied. An alternative method is the Frobenius norm, which was used by Filipiak et al. (2021), where the authors presented two methods for the identification of separable covariance structures with both components unstructured, or with one component additionally structured as CS or AR(1) for doubly multivariate data.

To confirm whether the choice of structure is appropriate, a hypothesis about the structure of the covariance matrix is tested, as a result of which we accept or reject the statistical hypothesis regarding the structure of the covariance matrix in the following form:

$$H_0 : \mathbf{\Omega} \quad \textit{is structured} \quad vs. \quad H_1 : \mathbf{\Omega} \quad \textit{is unstructured.}$$

Commonly used tests are LRT and RST, where the test statistics have the following forms:

$$\begin{aligned} LR &= n \operatorname{tr}\left(\widehat{\mathbf{\Omega}}_0 \widehat{\mathbf{\Omega}}_1^{-1}\right) - n \ln |\widehat{\mathbf{\Omega}}_0 \widehat{\mathbf{\Omega}}_1^{-1}| - nm \\ RS &= \tfrac{n}{2} \operatorname{tr}\left[(\widehat{\mathbf{\Omega}}_0 \widehat{\mathbf{\Omega}}_1^{-1} - \mathbf{I}_m)^2\right], \end{aligned}$$

where $\widehat{\mathbf{\Omega}}_0$ is an MLE of $\mathbf{\Omega}$ under the null hypothesis and $\widehat{\mathbf{\Omega}}_1$ is an MLE of $\mathbf{\Omega}$ under the alternative hypothesis. It is known that under the null hypothesis both of the tests have asymptotically χ^2 distribution with v degrees of freedom, where v is the difference between the number of unknown parameters under the alternative and null hypotheses (cf. Rao, 2005). The null hypothesis is rejected at the significance level α (e.g. 0.05 or 0.01) when $T > \kappa_\alpha$, where T is the LR or RS statistic and κ_α is the $(1-\alpha)100\%$ quantile of the distribution of the test statistic, if the null hypothesis is true (empirical null distribution). Filipiak and Klein (2021) compare both tests in the case of the covariance structure of double multivariate data. It was shown that, if the number of objects is small, RST converges faster to a χ^2 distribution. In addition, in a paper by Krishnaiah and Lee (1974), the authors consider the problem of testing the hypotheses that the covariance matrix of a multivariate normal population has a certain structure.

References

Anderson T. W. (1973): Asymptotically efficient estimation of covariance matrices with linear structure. Ann. Stat. 1(1): 135–141.

Andrews D. W. K., Ploberger, W. (1996): Testing for Serial Correlation Against an ARMA(1,1) Process. J. Am. Stat. Assoc, 91(435): 1331–1342.

Chakraborty M. (1998): An efficient algorithm for solving general periodic Toeplitz systems, in IEEE Transactions on Signal Processing, vol. 46, no. 3: 784-787.

Cui X., Lin C., Zhao J., Zeng L., Zhang D., Pan J. (2016): Covariance structure regularization via Frobenius norm discrepancy. Linear Algebra Appl. 510: 124–145.

Devijver E., Gallopin M. (2018): Block-Diagonal Covariance Selection for High-Dimensional Gaussian Graphical Models. J. Am. Stat. Assoc. 113(4): 306-314.

Filipiak K., John M., Markiewicz A. (2020): Comments on Maximum Likelihood Estimation and Projections Under Multivariate Statistical Models. In: Holgersson, T., Singull, M. (eds.) Recent Developments in Multivariate and Random Matrix Analysis, Springer, 51-66.

Filipiak K., Klein D. (2018): Approximation with a Kronecker product structure with one component as compound symmetry or autoregression. Electronic J. Linear Algebra. 559: 11-33.

Filipiak K., Klein D. (2021): Estimation and testing the covariance structure of doubly multivariate data. In: Filipiak, K., Markiewicz, A., von Rosen, D. (Eds.) Multivariate, Multilinear and Mixed Linear Models, Springer, 131-156.

Filipiak K., Klein D., Markiewicz A., Mokrzycka M. (2020): Approximation with a Kronecker product structure with one component as compound symmetry or autoregression via entropy loss function. Linear Algebra Appl. 610: 625-646.

Filipiak K., Klein D., Mokrzycka M. (2018a): Estimators comparison of separable covariance structure with one component as compound symmetry matrix. Electronic J. Linear Algebra. 33: 83-98.

Filipiak K., Klein D., Mokrzycka M. (2021): Separable covariance structure identification for doubly multivariate data. In: Filipiak, K., Markiewicz, A., von Rosen, D. (Eds.) Multivariate, Multilinear and Mixed Linear Models, Springer, 113-130.

Filipiak K., Markiewicz A., Mieldzioc A., Sawikowska A. (2018b): On projection of a positive definite matrix on a cone of nonnegative definite Toeplitz matrices. Electron. J. Linear Algebra. 33: 74-82.

Fuglede B, Jensen S. T. (2013): Positive projections of symmetric matrices and Jordan algebras. Expo. Math. 31: 295-303.

Gilson M., Dahmen D., Moreno-Bote R., Insabato A., Helias A. (2019): The covariance perceptron: A new framework for classification and processing of time series in recurrent neural networks. bioRxiv. https://doi.org/10.1101/562546.

Hannart A, Naveau P. (2014): Estimating high dimensional covariance matrices: A new look at the Gaussian conjugate framework. J Multivar Anal. 131: 149-162.

Ikeda Y, Kubokawa T, Srivastava MS.(2016): Comparison of linear shrinkage estimators of a large covariance matrix in normal and non-normal distributions. Comput Stat Data Anal. 95: 95-108.

Janiszewska M., Markiewicz A. (2022): Estimation of block covariance matrix with linearly structured blocks. Submitted.

Janiszewska M, Markiewicz A., Mokrzycka M. (2020): Block matrix approximation via entropy loss function. Appl. Math. 65 (6): 829-844.

Janiszewska M, Markiewicz A., Mokrzycka M. (2022): Quasi shrinkage estimation of block structured covariance matrix. J. Stat. Comput. Simul. Submitted.

John M., Mieldzioc A. (2020): The comparison of the estimators of banded Toeplitz covariance structure under the highdimensional multivariate model. Comm. Statist. Simulation Comput. 734–752.

Judge, G.G., Hill, R.C., Griffiths, W.E., Lutkepohl, H. ¨ and Lee, T.-C. (1982): Introduction to the theory and practice of econometrics, John Wiley & Sons: New York.

Krishnaiah P. R. and Lee Jack C. (1974): On Covariance Structures. Sankhya. 38 (4): 357-371.

Lee K, Baek C, Daniels MJ. 2017: ARMA Cholesky Factor Models for the Covariance Matrix of Linear Models. Comput Stat Data Anal. 115: 267-280.

Liang Y., Rosen von D., Rosen von T. (2020): On Properties of Toeplitz-type Covariance Matrices in Models with Nested Random Effects. Statistical Papers. 62 (6): 2509-2528.

Lin L., Higham N. J., Pan J. (2014): Covariance structure regularization via entropy loss function. Comput. Statist. Data Anal. 72: 315-327.

Lu N., Zimmerman D. L. (2005): The likelihood ratio test for a separable covariance matrix. Stat. Probab. Lett. 73: 449-457.

Markiewicz A, Mieldzioc A. (2022): Improved estimators of linearly structured covariance matrices. Submitted.

Mieldzioc A., Mokrzycka M., Sawikowska A. (2021): Identification of Block-Structured Covariance Matrix on an Example of Metabolomic Data. Separations. 8(11):205: 734-752.

Mokrzycka M. (2021): Aproksymacja macierzy kowariancji wybranymi strukturami w modelach podwójnie wielowymiarowych. [Doctoral dissertation thesis, Adam Mickiewicz University in Poznań, Poland].

Ning L., Jiang X., Gourgiou T. (2006): Geometric methods for estimation of structured covariances. Biometrica, 93(4): 927-941.

Pan J., Fang K. (2002): Growth Curve Models and Statistical Diagnostics. Springer-Verlag, New York.

Rao, C. R. (2005): Score test: historical review and recent developments. In Advances in Ranking and Selection, Multiple Comparisons, and Reliability, (Eds., N. Balakrishnan, N. Kannan, H.N. Nagaraja), Chapter 1, Birkhauser, Boston, 3-20.

Roy A., Zmyślony R., Fonseca M., Leiva R. (2016): Optimal estimation for doubly multivariate data in blocked compound symmetric covariance structure. J. Multivar. Anal. 144: 81–90.

Seely J. (1971): Quadratic Subspaces and Completeness. Ann. Statist. 42(2): 710-721.

Shi M., Xu Li., Sol'e P. (2021): On isodual double Toeplitz codes. Computer Science. ArXiv.

Srivastava M., Rosen von T, Rosen von D. (2008): Models with a Kronecker product covariance structure: Estimation and testing. Math. Methods Statist. 17: 357-370.

Szatrowski T. H. (1976): Estimation and testing for block compound symmetry and other patterned covariance matrices with linear and non-linear structure. Technical report No. 107, Dept. of Statistics, Stanford University.

Szatrowski T. H. (1982): Testing and estimation in the block compound symmetry problem. J. Educ. Stat. 7(1): 3-18.

Timm, N.H. (2002): Applied Multivariate Analysis. Springer-Verlag Inc., New York.

Tsay, S. R. (2002): Analysis of Financial Time Series. Canada: John Wiley & Sons Inc.

Woods J, (1972): Two-dimensional discrete Markovian fields IEEE Transactions on Information Theory. 18 (2): 232-240.

Wilks, S. S. (1946): Sample Criteria for Testing Equality of Means, Equality of Variances, and Equality of Covariances in a Normal Multivariate Distribution. Ann. Math. Stat. 17, no. 3: 257–81.