ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS DE MATHEMATICA
Volume 19, Number 1, June 2015
Available online at http://acutm.math.ut.ee

More on explicit estimators for a banded covariance matrix

Emil Karlsson and Martin Singull

Abstract. The problem of estimating mean and covariances of a multivariate normally distributed random vector has been studied in many forms. This paper focuses on the estimators proposed by Ohlson et al. (2011) for a banded covariance structure with m-dependence. We rewrite the estimator when $m = 1$, which makes it easier to analyze. This leads to an adjustment, and an unbiased estimator can be proposed. A new and easier proof of consistency is then presented.

This theory is also generalized to a general linear model where the corresponding theorems and propositions are stated to establish unbiasedness and consistency.

1. Introduction

There exist many estimates, tests, confidence intervals and types of regression models in the multivariate statistical literature that are based on the assumption that the underlying distribution is normal [1, 9, 15]. The primary reason is that often multivariate datasets are, at least approximately, normally distributed. The multivariate normal distribution is also simpler to analyze than many other distributions. For example all the information in a multivariate normal distribution can be found in its mean and covariances. Because of this, estimating the mean and covariances are subjects of importance in statistics.

This paper will study an estimating procedure of a patterned covariance matrix. Patterned covariance matrices arise from a variety of different situations and applications and have been studied by many authors. In a seminal paper in the 1940s, Wilks [17] considered patterned covariances when studying psychological tests. Wilks [17] used the covariance matrix with equal

Received April 24, 2015.

2010 *Mathematics Subject Classification.* 62H12.

Key words and phrases. Banded covariance matrices, covariance matrix estimation, explicit estimators, multivariate normal distribution, general linear model.

http://dx.doi.org/10.12697/ACUTM.2015.19.05

diagonal and equal off-diagonal elements, called the *intraclass* covariance structure. Two years later Votaw [16] extended the intraclass covariance structure to a model with blocks which had a certain pattern, the so-called *compound symmetry* of *type I* and *type II*.

Olkin and Press [14] considered three symmetries, namely *circular*, intraclass and spherical and derived likelihood ratio test and the asymptotic distribution under the hypothesis and alternative. Olkin [13] generalized the circular stationary model with a multivariate version in which each element was a vector and the covariance matrix can be written as a *block circular* matrix.

The covariance symmetries investigated, for example, in [17, 16] and [14] are all special cases of *invariant normal models* considered by [2].

Permutation invariant covariance matrices were considered in [10] and it was proven that permutation invariance implies a specific structure for the covariance matrix. Nahtman and von Rosen [11] showed that shift invariance implies *Toeplitz* covariance matrices and marginally shift invariance gives *block Toeplitz* covariance matrices.

There exist many papers on Toeplitz covariance matrices, e.g., see [3], [6], [7] and [5]. To have a Toeplitz structure means that certain invariance conditions are fulfilled, e.g., equality of variances and covariances. A similar structure as the Topelitz structure is the *banded covariance matrix*. Banded covariance matrices are common in applications and arise often in association with time series. For example in signal processing, covariances of Gauss–Markov random processes or cyclostationary processes [18, 8, 4]. In this paper we will study a special case of banded matrices with unequal elements except that certain covariances are zero. These covariance matrices will have a tridiagonal structure.

Originally, estimates of covariance matrices were obtained using non-iterative methods such as analysis of variance and minimum norm quadratic unbiased estimation. Modern computers have changed a lot of things and the cheap processing power made it possible to use iterative methods which perform better. With this came the rise of the maximum likelihood method and more general estimating equations. These methods have surely dominated during the last years. But nowadays we see a shift back to non-iterative methods since the datasets have grown tremendously. With huge datasets estimating with iterative methods can be a slow and tedious job.

This paper will discuss some properties of an explicit non-iterative estimator for a banded covariance matrix derived in [12] and present an improvement to this estimator. The improvement gives an unbiased and consistent estimator for the mean and the covariance matrix under the special case of first order dependence.

The outline of the paper is as follows. Section 2 presents the explicit estimator given by [12] and some results regarding it. From these results a new unbiased explicit estimator is suggested. In Section 3 the explicit estimator is generalized for estimating the covariance matrix in a general linear model. An unbiased estimator is proposed. We conclude with a small simulation study in Section 4 which is based on the new unbiased explicit estimator proposed in this paper and some conclusions in Section 5.

2. Explicit estimators of a banded covariance matrix

In [12] an explicit estimator for the covariance matrix for a multivariate normal distribution when the covariance matrix have an m-dependence structure is presented. Ohlson et al. [12] propose estimators for the general case when $m+1<p<n$ and establish some properties of it. Furthermore, they also consider the special case where $m=1$ in detail and in this section some of the results from the article will be presented. The banded covariance structure of order one is given by

$$\boldsymbol{\Sigma}_{(p)}^{(1)} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & 0 & \dots & \dots & 0 \\ \sigma_{12} & \sigma_{22} & \sigma_{23} & 0 & \dots & 0 \\ 0 & \sigma_{32} & \sigma_{33} & \sigma_{34} & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \sigma_{p-2,p-1} & \sigma_{p-1,p-1} & \sigma_{p-1,p} \\ 0 & \dots & \dots & 0 & \sigma_{p-1,p} & \sigma_{pp} \end{pmatrix}. \tag{1}$$

Given the observation matrix

$$\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{y}_1 & \dots & \boldsymbol{y}_p \end{pmatrix}' \sim N_{p,n}\left(\boldsymbol{\mu}\mathbf{1}_n', \boldsymbol{\Sigma}_{(p)}^{(1)}, \boldsymbol{I}_n\right), \tag{2}$$

where $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 & \dots & \mu_p \end{pmatrix}'$ and $\boldsymbol{I}_n$ is the identity matrix of order n, the estimators $\hat{\sigma}_{ii}$ and $\hat{\sigma}_{i,i+1}$ are constructed through conditioning on $\boldsymbol{y}_1, \dots, \boldsymbol{y}_{i-1}$.

2.1. Previous results. Below follows the proposition given in [12].

Proposition 2.1. *Let $\boldsymbol{Y} \sim N_{p,n}\left(\boldsymbol{\mu}\mathbf{1}_n', \boldsymbol{\Sigma}_{(p)}^{(1)}, \boldsymbol{I}_n\right)$. Explicit estimators are given by*

$$\begin{aligned} \hat{\mu}_i &= \frac{1}{n}\boldsymbol{y}_i'\mathbf{1}_n, \\ \hat{\sigma}_{ii} &= \frac{1}{n}\boldsymbol{y}_i'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_i, \qquad \textit{for } i=1,\dots,p, \\ \hat{\sigma}_{i,i+1} &= \frac{1}{n}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_{i+1}, \qquad \textit{for } i=1,\dots,p-1, \end{aligned}$$

where $\hat{\boldsymbol{r}}_1 = \boldsymbol{y}_1$, $\hat{\boldsymbol{r}}_i = \boldsymbol{y}_i - \hat{s}_i\hat{\boldsymbol{r}}_{i-1}$ *for* $i = 2, \ldots, p-1$, *and* $\hat{s}_i = \dfrac{\hat{\boldsymbol{r}}'_{i-1}\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_i}{\hat{\boldsymbol{r}}'_{i-1}\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_{i-1}}$, *where*

$$\boldsymbol{Q}_{\mathbf{1}_n} = \boldsymbol{I}_n - \mathbf{1}_n(\mathbf{1}'_n\mathbf{1}_n)^{-1}\mathbf{1}'_n \tag{3}$$

and $\mathbf{1} = \begin{pmatrix}1 & \ldots & 1\end{pmatrix}' : n \times 1$.

Some properties of the estimators where given in [12], where it was shown that the estimator $\hat{\boldsymbol{\Sigma}}^{(1)}_{(p)} = (\hat{\sigma}_{ij})$ given in Proposition 2.1 is consistent. But, the estimator for the covariance matrix above lacks the property of unbiasedness. One of the main goals of this paper is to develop this desired property.

2.2. Remodeling of explicit estimators. We will now rewrite the estimators given in Proposition 2.1 which makes it clearer and more suitable for interpretation and analyzes.

The estimators presented in Proposition 2.1 are partly composed from the maximum likelihood estimators (MLEs). The estimator for μ_i, given by

$$\hat{\mu}_i = \frac{1}{n}\boldsymbol{y}'_i\mathbf{1}_n,$$

is the MLE for the unstructured case by construction.

The proposed estimator and the MLEs for an unstructured covariance matrix share a resemblance, which can be seen by looking at the estimators of the diagonal elements, which are given as

$$\hat{\sigma}_{ii} = \frac{1}{n}\boldsymbol{y}'_i\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_i, \qquad \text{for } i = 1, \ldots, p,$$

where $\boldsymbol{Q}_{\mathbf{1}_n}$ is given in (3). These are the same for the two cases. Also the first off-diagonal element is of course the same as the MLE for the unstructed case, since this is how it is constructed. The estimator is given by

$$\hat{\sigma}_{12} = \frac{1}{n}\hat{\boldsymbol{r}}'_1\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_2,$$

where

$$\hat{\boldsymbol{r}}_1 = \boldsymbol{y}_1.$$

However, the estimators of the off-diagonal elements of the covariance matrix except σ_{12} are not the same as the MLEs for the unstructured covariance matrix and are therefore not so straightforward to analyze.

Furthermore, when $i > 1$ we can write the estimators as

$$\begin{aligned}
\hat{\sigma}_{i,i+1} &= \frac{1}{n}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_{i+1} = \frac{1}{n}(\boldsymbol{y}_i - \hat{s}_i\hat{\boldsymbol{r}}_{i-1})'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_{i+1} \\
&= \frac{1}{n}\left(\boldsymbol{y}_i - \frac{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_i}{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_{i-1}}\hat{\boldsymbol{r}}_{i-1}\right)'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_{i+1} \\
&= \frac{1}{n}\left(\boldsymbol{y}_i'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_{i+1} - \frac{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_i}{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_{i-1}}\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_{i+1}\right)
\end{aligned}$$

and since $\hat{\boldsymbol{r}}_{k-1}'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_{k-1}$ is a scalar it is possible to write

$$\begin{aligned}
\hat{\sigma}_{i,i+1} &= \frac{1}{n}\left(\boldsymbol{y}_i'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_{i+1} - \boldsymbol{y}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_{i-1}(\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_{i-1})^{-1}\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_{i+1}\right) \\
&= \frac{1}{n}\boldsymbol{y}_i'\left(\boldsymbol{Q}_{\mathbf{1}_n} - \boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_{i-1}(\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_{i-1})^{-1}\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\mathbf{1}_n}\right)\boldsymbol{y}_{i+1} \\
&= \frac{1}{n}\boldsymbol{y}_i'\boldsymbol{P}_{\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i}\boldsymbol{y}_{i+1},
\end{aligned}$$

where

$$\boldsymbol{P}_{\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i} = \boldsymbol{Q}_{\mathbf{1}_n} - \boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}.$$

For simplicity we will write

$$\boldsymbol{P}_i^{\mathbf{1}} = \boldsymbol{P}_{\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i}. \tag{4}$$

The main proposition of this paper follows, i.e., an alternative writing of Proposition 2.1.

Proposition 2.2. *Let* $\boldsymbol{Y} \sim N_{p,n}\left(\boldsymbol{\mu}\mathbf{1}_n', \boldsymbol{\Sigma}_{(p)}^{(1)}, \boldsymbol{I}_n\right)$. *Explicit estimators of the parameters are given by*

$$\begin{aligned}
\hat{\mu}_i &= \frac{1}{n}\boldsymbol{y}_i'\mathbf{1}_n, \\
\hat{\sigma}_{ii} &= \frac{1}{n}\boldsymbol{y}_i'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_i, \qquad \textit{for } i = 1, \dots, p, \\
\hat{\sigma}_{i,i+1} &= \frac{1}{n}\boldsymbol{y}_i'\boldsymbol{P}_{i-1}^{\mathbf{1}}\boldsymbol{y}_{i+1}, \qquad \textit{for } i = 1, \dots, p-1,
\end{aligned}$$

where $\boldsymbol{P}_i^{\mathbf{1}} = \boldsymbol{Q}_{\mathbf{1}_n} - \boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}$ *with* $\hat{\boldsymbol{r}}_1 = \mathbf{0}$ *and*

$$\hat{\boldsymbol{r}}_i = \boldsymbol{y}_i - \frac{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_i}{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_{i-1}}\hat{\boldsymbol{r}}_{i-1}, \qquad \textit{for } i = 2, \dots, p-1,$$

and $\boldsymbol{Q}_{\mathbf{1}_n}$ *is given in* (3).

The next theorem shows an important property of the matrix $\boldsymbol{P}_i^{\mathbf{1}}$.

Theorem 2.1. *The matrix* $\boldsymbol{P}_i^{\mathbf{1}}$, $i = 2, \dots, p-2$, *used in Proposition 2.2 is idempotent and symmetric of rank* $n - 2$.

14

Proof. Idempotency: We have

$$\begin{aligned}\boldsymbol{P}_i^{\mathbf{1}^2} &= (\boldsymbol{Q}_{\mathbf{1}_n} - \boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n})(\boldsymbol{Q}_{\mathbf{1}_n} - \boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n})\\ &= \boldsymbol{Q}_{\mathbf{1}_n}^2 - \boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n} - \boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{Q}_{\mathbf{1}_n}\\ &\qquad + \boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\\ &= \boldsymbol{Q}_{\mathbf{1}_n} - \boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n} = \boldsymbol{P}_i^{\mathbf{1}},\end{aligned}$$

since $\boldsymbol{Q}_{\mathbf{1}_n}$, given in (3), is an idempotent matrix.

Symmetry: $\boldsymbol{P}_i$ is symmetric since

$$\begin{aligned}\boldsymbol{P}_i^{\mathbf{1}'} &= (\boldsymbol{Q}_{\mathbf{1}_n} - \boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n})' = \boldsymbol{Q}_{\mathbf{1}_n}' - \boldsymbol{Q}_{\mathbf{1}_n}'\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}'\\ &= \boldsymbol{Q}_{\mathbf{1}_n} - \boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n} = \boldsymbol{P}_i^{\mathbf{1}}.\end{aligned}$$

Rank: Since $\boldsymbol{P}_i^{\mathbf{1}}$ is idempotent, the $\mathrm{rank}(\boldsymbol{P}_i^{\mathbf{1}}) = \mathrm{tr}(\boldsymbol{P}_i^{\mathbf{1}})$. This implies the following:

$$\begin{aligned}\mathrm{rank}(\boldsymbol{P}_i^{\mathbf{1}}) = \mathrm{tr}(\boldsymbol{P}_i^{\mathbf{1}}) &= \mathrm{tr}(\boldsymbol{Q}_{\mathbf{1}_n} - \boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n})\\ &= \mathrm{tr}(\boldsymbol{Q}_{\mathbf{1}_n}) - \mathrm{tr}(\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n})\\ &= n-1-\mathrm{tr}((\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}) = n-2.\end{aligned}$$

□

2.3. Unbiasedness and consistency. The last section presented some alteration to the original estimators which made it possible to rewrite them as a quadratic and bilinear forms, centering with an idempotent matrix, i.e.,

$$\hat{\sigma}_{ii} = \frac{1}{n}\boldsymbol{y}_i'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_i, \qquad \text{for } i = 1,\dots,p,$$

and

$$\hat{\sigma}_{i,i+1} = \frac{1}{n}\boldsymbol{y}_i'\boldsymbol{P}_{i-1}^{\mathbf{1}}\boldsymbol{y}_{i+1}, \qquad \text{for } i = 1,\dots,p-1.$$

We can now propose an unbiased estimator for the covariance matrix. It is also possible to present a new and much simpler proof of the consistency for the sample covariance matrix compared to the proof given in [12].

First we propose an unbiased estimator for the covariance matrix.

Theorem 2.2. *Let* $\boldsymbol{Y} \sim N_{p,n}\left(\boldsymbol{\mu}\mathbf{1}_n', \boldsymbol{\Sigma}_{(p)}^{(1)}, \boldsymbol{I}_n\right)$. *Explicit unbiased estimators of the parameters are given by*

$$\begin{aligned}
\hat{\mu}_i &= \frac{1}{n}\boldsymbol{y}_i'\mathbf{1}_n, \\
\hat{\sigma}_{ii} &= \frac{1}{n-1}\boldsymbol{y}_i'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_i, \qquad \text{for } i = 1,\ldots,p, \\
\hat{\sigma}_{12} &= \frac{1}{n-1}\boldsymbol{y}_1'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_2, \\
\hat{\sigma}_{i,i+1} &= \frac{1}{n-2}\boldsymbol{y}_i'\boldsymbol{P}_{i-1}^{\mathbf{1}}\boldsymbol{y}_{i+1}, \qquad \text{for } i = 2,\ldots,p-1,
\end{aligned}$$

where $\boldsymbol{P}_i^{\mathbf{1}} = \boldsymbol{Q}_{\mathbf{1}_n} - \boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\mathbf{1}_n}$ *with* $\hat{\boldsymbol{r}}_1 = \boldsymbol{y}_1$ *and*

$$\hat{\boldsymbol{r}}_i = \boldsymbol{y}_i - \frac{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_i}{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\mathbf{1}_n}\boldsymbol{y}_{i-1}}\hat{\boldsymbol{r}}_{i-1}, \qquad \text{for } i = 2,\ldots,p-1,$$

and $\boldsymbol{Q}_{\mathbf{1}_n}$ *is given in* (3).

Proof. The estimators $\hat{\mu}_i$, $\hat{\sigma}_{ii}$ for $i = 1,\ldots,p$ and $\hat{\sigma}_{12}$ coincide with the corrected maximum likelihood estimators and are thus unbiased. Therefore it remains to prove that $\hat{\sigma}_{i,i+1}$ are unbiased for $i = 2,\ldots,p-1$.

In the derivation of $\hat{\sigma}_{i,i+1}$ we assume $\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{i-1}$ to be known, see [12] for more details. Therefore, the matrix $\boldsymbol{P}_i^{\mathbf{1}}$, $i = 2,\ldots,p-2$, can be considered as a non-random matrix. We consider $\hat{\sigma}_{i,i+1}$ as a bilinear form and calculate its expected value as

$$\begin{aligned}
\mathrm{E}(\hat{\sigma}_{i,i+1}) &= \mathrm{E}\left(\frac{1}{n-2}\boldsymbol{y}_i'\boldsymbol{P}_{i-1}^{\mathbf{1}}\boldsymbol{y}_{i+1}\right) = \frac{1}{n-2}\,\mathrm{E}\left[\mathrm{E}(\boldsymbol{y}_i'\boldsymbol{P}_{i-1}^{\mathbf{1}}\boldsymbol{y}_{i+1}|\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{i-1})\right] \\
&= \frac{1}{n-2}\,\mathrm{E}\left[\mathrm{tr}(\boldsymbol{P}_{i-1}^{\mathbf{1}})\right]\sigma_{i,i+1} = \sigma_{i,i+1},
\end{aligned}$$

since the matrix $\boldsymbol{P}_{i-1}^{\mathbf{1}}$ is idempotent we have $\mathrm{tr}(\boldsymbol{P}_{i-1}^{\mathbf{1}}) = \mathrm{rank}(\boldsymbol{P}_{i-1}^{\mathbf{1}}) = n-2$, and the theorem has been proved.

□

Theorem 2.3. *The estimators given in Theorem 2.2 are consistent.*

The proof for consistency follows the same idea as the proof for unbiasedness.

Proof. The estimators $\hat{\mu}_i$, $\hat{\sigma}_{ii}$ for $i = 1,\ldots,p$ and $\hat{\sigma}_{12}$ coincide with the corrected maximum likelihood estimators and are thus consistent. Therefore it remains to prove that $\hat{\sigma}_{i,i+1}$ are consistent for $i = 2,\ldots,p-1$.

We consider $\hat{\sigma}_{i,i+1}$ as a bilinear form and calculate its variance as

$$\begin{aligned}\operatorname{var}(\hat{\sigma}_{i,i+1}) &= \operatorname{var}\left(\frac{1}{n-2}\boldsymbol{y}_i'\boldsymbol{P}^{\mathbf{1}}_{i-1}\boldsymbol{y}_{i+1}\right)\\ &= \frac{1}{(n-2)^2}\Big(\mathrm{E}\left[\operatorname{var}(\boldsymbol{y}_i'\boldsymbol{P}^{\mathbf{1}}_{i-1}\boldsymbol{y}_{i+1}|\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{i-1})\right]\\ &\qquad\qquad + \underbrace{\operatorname{var}\left[\mathrm{E}(\boldsymbol{y}_i'\boldsymbol{P}^{\mathbf{1}}_{i-1}\boldsymbol{y}_{i+1}|\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{i-1})\right]}_{=0}\Big)\\ &= \frac{1}{(n-2)^2}\,\mathrm{E}\left[\operatorname{tr}(\boldsymbol{P}^{\mathbf{1}}_{i-1})\right]\sigma^2_{i,i+1} + \mathrm{E}\left[\operatorname{tr}({\boldsymbol{P}^{\mathbf{1}}_{i-1}}^2)\right]\sigma_{ii}\sigma_{i+1,i+1}.\end{aligned}$$

Since the matrix $\boldsymbol{P}^{\mathbf{1}}_{i-1}$ is idempotent, we have

$$\begin{aligned}\operatorname{var}(\hat{\sigma}_{i,i+1}) &= \frac{1}{(n-2)^2}\operatorname{rank}(\boldsymbol{P}^{\mathbf{1}}_{i-1})(\sigma^2_{i,i+1}+\sigma_{ii}\sigma_{i+1,i+1})\\ &= \frac{\sigma^2_{i,i+1}+\sigma_{ii}\sigma_{i+1,i+1}}{n-2},\end{aligned}$$

since $\operatorname{rank}(\boldsymbol{P}^{\mathbf{1}}_{i-1}) = n-2$. Hence, $\operatorname{var}(\hat{\sigma}_{i,i+1}) \to 0$, when $n \to \infty$. The estimator is unbiased, hence the consistency follows. Thus the theorem has been proved. □

3. Generalization to a general linear model

In this section the estimator presented earlier will be extended to a general linear model. Two differences of concern are the effect of estimating the regression parameters and the degrees of freedom, i.e., the rank of the design matrix. The multivariate linear model takes the form

$$\boldsymbol{Y} = \boldsymbol{B}\boldsymbol{X} + \boldsymbol{E} : p \times n,$$

where $\boldsymbol{X} : k \times n$ is a known design matrix, and

$$\boldsymbol{B} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_p)' : p \times k$$

is an unknown matrix of regression parameters. We will assume throughout this paper, without loss of generality, that $\boldsymbol{X}$ has full rank k, such that $n \geq p + k$, and the error matrix is normally distributed, i.e., $\boldsymbol{E} \sim N_{p,n}(\mathbf{0}, \boldsymbol{\Sigma}, \boldsymbol{I}_n)$.

3.1. $\hat{\boldsymbol{B}}$ instead of $\hat{\boldsymbol{\mu}}$. This section contains a motivation why $\hat{\boldsymbol{B}}$ will maximize the conditional likelihood function in the same way as $\hat{\boldsymbol{\mu}}$ does.

In a general linear model the MLE for $\boldsymbol{B}$ is $\hat{\boldsymbol{B}} = \boldsymbol{Y}\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{X}')^{-1}$ for the unstructured case. Since the general linear model is a fusion between different response value, it is possible to determine the different rows $\boldsymbol{b}_i'$ separately with the following expression

$$\hat{\boldsymbol{b}}_i' = \boldsymbol{y}_i'\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{X}')^{-1}, \quad \text{for } i = 1, \ldots, k.$$

The explicit estimator in [12] is derived from a stepwise maximization of the likelihood function. The same principle applies for the general linear model in the following way

$$\hat{\boldsymbol{b}}_i' = \hat{\boldsymbol{b}}_i' | \boldsymbol{y}_1', \ldots, \boldsymbol{y}_{i-1}'.$$

Since each individual $\boldsymbol{b}$-vector can be determined independently and because the estimator of $\hat{\boldsymbol{b}}$ above is the MLE for the unstructured case, it will maximize the conditional distribution and we have

$$\hat{\boldsymbol{B}} = (\hat{\boldsymbol{b}}_1, \ldots, \hat{\boldsymbol{b}}_p)'.$$

Altogether this makes a good basis to propose explicit estimators for a general linear model with banded structure of order one.

3.2. Proposed estimators. In this section we propose explicit estimators for a general linear model. In the section above we motivated the estimator $\hat{\boldsymbol{B}} = \boldsymbol{Y}\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{X}')^{-1}$ and here follows a proposition for the covariance matrix. In Section 2 we assumed $\boldsymbol{Y} \sim N_{p,n}(\boldsymbol{\mu}\mathbf{1}_n', \boldsymbol{\Sigma}_{(p)}^{(1)}, \boldsymbol{I}_n)$. We now study the general linear model $\boldsymbol{Y} \sim N_{p,n}(\boldsymbol{B}\boldsymbol{X}, \boldsymbol{\Sigma}_{(p)}^{(1)}, \boldsymbol{I}_n)$ and see that the transformation $\boldsymbol{Y} - \boldsymbol{B}\boldsymbol{X}$ will yield the same model as in Section 2, i.e., $\boldsymbol{Y} - \boldsymbol{B}\boldsymbol{X} \sim N_{p,n}(\mathbf{0}, \boldsymbol{\Sigma}_{(p)}^{(1)}, \boldsymbol{I}_n)$. Hence, in Theorem 2.2 we will now replace $\boldsymbol{y}_i'$ with

$$\boldsymbol{y}_i' - \hat{\boldsymbol{b}}_i'\boldsymbol{X} = \boldsymbol{y}_i'(\boldsymbol{I}_n - \boldsymbol{X}'(\boldsymbol{X}\boldsymbol{X}')^{-1}\boldsymbol{X}) = \boldsymbol{y}_i'\boldsymbol{Q}_{\boldsymbol{X}},$$

where

$$\boldsymbol{Q}_{\boldsymbol{X}} = \boldsymbol{I}_n - \boldsymbol{X}'(\boldsymbol{X}\boldsymbol{X}')^{-1}\boldsymbol{X}. \tag{5}$$

This leads us to the following proposition.

Proposition 3.1. *Let* $\boldsymbol{Y} \sim N_{p,n}\left(\boldsymbol{B}\boldsymbol{X}, \boldsymbol{\Sigma}_{(p)}^{(1)}, \boldsymbol{I}_n\right)$, *with* $\operatorname{rank}(\boldsymbol{X}) = k$. *Explicit estimators of the parameters are given by*

$$\begin{aligned}
\hat{\boldsymbol{B}} &= \boldsymbol{Y}\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{X}')^{-1}, \\
\hat{\sigma}_{ii} &= \frac{1}{n-1}\boldsymbol{y}_i'\boldsymbol{Q}_{\boldsymbol{X}}\boldsymbol{y}_i, \qquad \textit{for } i = 1, \ldots, p, \\
\hat{\sigma}_{12} &= \frac{1}{n-2}\boldsymbol{y}_1'\boldsymbol{Q}_{\boldsymbol{X}}\boldsymbol{y}_2, \\
\hat{\sigma}_{i,i+1} &= \frac{1}{n-2}\boldsymbol{y}_i'\boldsymbol{P}_{i-1}^{\boldsymbol{X}}\boldsymbol{y}_{i+1}, \qquad \textit{for } i = 2, \ldots, p-1,
\end{aligned}$$

where $\boldsymbol{P}_i^{\boldsymbol{X}} = \boldsymbol{Q}_{\boldsymbol{X}} - \boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}$ *with* $\hat{\boldsymbol{r}}_1 = \boldsymbol{y}_1$ *and*

$$\hat{\boldsymbol{r}}_i = \boldsymbol{y}_i - \frac{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\boldsymbol{X}}\boldsymbol{y}_i}{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\boldsymbol{X}}\boldsymbol{y}_{i-1}}\hat{\boldsymbol{r}}_{i-1}, \qquad \textit{for } i = 2, \ldots, p-1$$

and $\boldsymbol{Q}_{\boldsymbol{X}}$ *is given in* (5).

15

In Section 2 we saw that the correction of unbiasedness depended on matrix $\boldsymbol{P}_i^{\mathbf{1}}$ which $\boldsymbol{Q}_{\mathbf{1}_n}$ is a part of, we need to study the properties of the new matrix $\boldsymbol{P}_i^{\boldsymbol{X}}$ to determine what kind of estimator for a general linear model will give us unbiasedness.

Here follows a theorem regarding the properties of the matrix $\boldsymbol{P}_i^{\boldsymbol{X}}$ above.

Theorem 3.1. *The matrix $\boldsymbol{P}_i^{\boldsymbol{X}}$ given in Proposition 3.1 is idempotent and symmetric with* $\operatorname{rank}(\boldsymbol{P}_i^{\boldsymbol{X}}) = n-k-1$.

Proof. Idempotence:

$$\begin{aligned}
{\boldsymbol{P}_i^{\boldsymbol{X}}}^2 &= (\boldsymbol{Q}_{\boldsymbol{X}} - \boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}})(\boldsymbol{Q}_{\boldsymbol{X}} - \boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}) \\
&= \boldsymbol{Q}_{\boldsymbol{X}}^2 - \boldsymbol{Q}_{\boldsymbol{X}}\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}} - \boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\boldsymbol{Q}_{\boldsymbol{X}} \\
&\qquad + \boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}} \\
&= \boldsymbol{Q}_{\boldsymbol{X}} - \boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}} = \boldsymbol{P}_i^{\boldsymbol{X}},
\end{aligned}$$

since $\boldsymbol{Q}_{\boldsymbol{X}}$ is an idempotent matrix.

Symmetry:

$$\begin{aligned}
{\boldsymbol{P}_i^{\boldsymbol{X}}}' &= (\boldsymbol{Q}_{\boldsymbol{X}} - \boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}})' \\
&= \boldsymbol{Q}_{\boldsymbol{X}}' - \boldsymbol{Q}_{\boldsymbol{X}}'\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}' \\
&= \boldsymbol{Q}_{\boldsymbol{X}} - \boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}} = \boldsymbol{P}_i^{\boldsymbol{X}}.
\end{aligned}$$

Rank: Since $\boldsymbol{P}_i^{\boldsymbol{X}}$ is idempotent, the $\operatorname{rank}(\boldsymbol{P}_i^{\boldsymbol{X}}) = \operatorname{tr}(\boldsymbol{P}_i^{\boldsymbol{X}})$. This implies

$$\begin{aligned}
\operatorname{rank}(\boldsymbol{P}_i^{\boldsymbol{X}}) &= \operatorname{tr}(\boldsymbol{P}_i^{\boldsymbol{X}}) = \operatorname{tr}(\boldsymbol{Q}_{\boldsymbol{X}} - \boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}) \\
&= \operatorname{tr}(\boldsymbol{Q}_{\boldsymbol{X}}) - \operatorname{tr}(\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}) \\
&= n - k - \operatorname{tr}((\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}) = n - k - 1.
\end{aligned}$$

□

3.3. Unbiasedness and consistency. Given Theorem 3.1 we are now again ready to propose an unbiased estimator. Since the structure of the estimators is similar to the multivariate normal model discussed in Section 2, the proofs will be similar.

Theorem 3.2. *Let* $\boldsymbol{Y} \sim N_{p,n}\left(\boldsymbol{BX}, \boldsymbol{\Sigma}_{(p)}^{(1)}, \boldsymbol{I}_n\right)$*, where* $\text{rank}(\boldsymbol{X}) = k$*. Explicit unbiased estimators of the parameters are given by*

$$\hat{\boldsymbol{B}} = \boldsymbol{YX}'(\boldsymbol{XX}')^{-1},$$

$$\hat{\sigma}_{ii} = \frac{1}{n-k}\boldsymbol{y}_i'\boldsymbol{Q}_{\boldsymbol{X}}\boldsymbol{y}_i, \qquad \text{for } i = 1,\dots,p,$$

$$\hat{\sigma}_{12} = \frac{1}{n-k}\boldsymbol{y}_1'\boldsymbol{Q}_{\boldsymbol{X}}\boldsymbol{y}_2,$$

$$\hat{\sigma}_{i,i+1} = \frac{1}{n-k-1}\boldsymbol{y}_i'\boldsymbol{P}_{i-1}^{\boldsymbol{X}}\boldsymbol{y}_{i+1}, \qquad \text{for } i = 2,\dots,p-1,$$

where $\boldsymbol{P}_i^{\boldsymbol{X}} = \boldsymbol{Q}_{\boldsymbol{X}} - \boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{Q}_{\boldsymbol{X}}$ *with* $\hat{\boldsymbol{r}}_1 = \boldsymbol{y}_1$ *and*

$$\hat{\boldsymbol{r}}_i = \boldsymbol{y}_i - \frac{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\boldsymbol{X}}\boldsymbol{y}_i}{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{Q}_{\boldsymbol{X}}\boldsymbol{y}_{i-1}}\hat{\boldsymbol{r}}_{i-1}, \qquad \text{for } i = 2,\dots,p-1,$$

and $\boldsymbol{Q}_{\boldsymbol{X}}$ *is given in* (5).

Proof. The estimators $\hat{\boldsymbol{B}}$, $\hat{\sigma}_{ii}$ for $i = 1,\dots,p$ and $\hat{\sigma}_{12}$ coincide with the corrected maximum likelihood estimators and are thus unbiased. It remains to prove that $\hat{\sigma}_{i,i+1}$ are unbiased for $i = 2,\dots,p-1$.

When the estimators are derived, they are conditioned on the previous $\boldsymbol{y}$:s. That is the calculation of $\hat{\sigma}_{i,i+1}$ assumes $\boldsymbol{y}_1,\dots,\boldsymbol{y}_{i-1}$ to be known constants. Therefore, the matrix $\boldsymbol{P}_i^{\boldsymbol{X}}$ below can be considered as a non-random matrix.

We can consider $\hat{\sigma}_{i,i+1}$ as a bilinear form and calculate its expectation as

$$\begin{aligned}
\mathrm{E}(\hat{\sigma}_{i,i+1}) &= \mathrm{E}\left(\frac{1}{n-k-1}\boldsymbol{y}_i'\boldsymbol{P}_{i-1}^{\boldsymbol{X}}\boldsymbol{y}_{i+1}\right)\\
&= \frac{1}{n-k-1}\,\mathrm{E}\left[\mathrm{E}\left(\boldsymbol{y}_i'\boldsymbol{P}_{i-1}^{\boldsymbol{X}}\boldsymbol{y}_{i+1}|\boldsymbol{y}_1,\dots,\boldsymbol{y}_{i-1}\right)\right]\\
&= \frac{1}{n-k-1}\,\mathrm{E}\left[\mathrm{tr}(\boldsymbol{P}_{i-1}^{\boldsymbol{X}})\right]\sigma_{i,i+1} = \sigma_{i,i+1},
\end{aligned}$$

since $\mathrm{tr}(\boldsymbol{P}_{i-1}^{\boldsymbol{X}}) = \mathrm{rank}(\boldsymbol{P}_{i-1}^{\boldsymbol{X}}) = n-k-1$. Thus $\mathrm{E}(\hat{\sigma}_{i,i+1}) = \sigma_{i,i+1}$ and the theorem has been proved. □

The proof for consistency follows the same structure as the proof above but instead uses that the estimators are unbiased and study the variance of the estimators.

Theorem 3.3. *The estimators given in Theorem 3.2 are consistent.*

Proof. The estimators $\hat{\mu}_i$, $\hat{\sigma}_{ii}$ for $i = 1,\dots,p$ and $\hat{\sigma}_{12}$ coincide with the corrected maximum likelihood estimators and are thus consistent.

It remains to prove that $\hat{\sigma}_{i,i+1}$ are consistent for $i = 2,\dots,p-1$. When the estimators are derived, they are conditioned on the previous $\boldsymbol{y}$:s. That is the calculation of $\hat{\sigma}_{i,i+1}$ assumes $\boldsymbol{y}_1,\dots,\boldsymbol{y}_{i-1}$ to be known constants. Therefore, the matrix $\boldsymbol{P}_i^{\boldsymbol{X}}$ below can be considered as a non-random matrix. We can then consider $\hat{\sigma}_{i,i+1}$ as a bilinear form and calculate its variance as

$$
\begin{aligned}
\operatorname{var}(\hat{\sigma}_{i,i+1}) &= \operatorname{var}\left(\frac{1}{n-k-1}\boldsymbol{y}_i'\boldsymbol{P}_{i-1}^{\boldsymbol{X}}\boldsymbol{y}_{i+1}\right) \\
&= \frac{1}{(n-k-1)^2}\,\mathrm{E}\left[\operatorname{var}(\boldsymbol{y}_i'\boldsymbol{P}_{i-1}^{\boldsymbol{X}}\boldsymbol{y}_{i+1}|\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{i-1})\right] \\
&= \frac{1}{(n-k-1)^2}\left(\mathrm{E}\left[\operatorname{tr}(\boldsymbol{P}_{i-1}^{\boldsymbol{X}})\right]\sigma_{i,i+1}^2 + \mathrm{E}\left[\operatorname{tr}({\boldsymbol{P}_{i-1}^{\boldsymbol{X}}}^2)\right]\sigma_{ii}\sigma_{i+1,i+1}\right) \\
&= \frac{\sigma_{i,i+1}^2 + \sigma_{ii}\sigma_{i+1,i+1}}{n-k-1},
\end{aligned}
$$

since the matrix $\boldsymbol{P}_{i-1}^{\boldsymbol{X}}$ is idempotent and $\operatorname{tr}(\boldsymbol{P}_{i-1}^{\boldsymbol{X}}) = \operatorname{rank}(\boldsymbol{P}_{i-1}^{\boldsymbol{X}}) = n-k-1$. One can see that, when $n \to \infty$ the $\operatorname{var}(\hat{\sigma}_{i,i+1}) \to 0$, i.e., consistent. Thus the theorem has been proved. $\square$

4. Simulations

In this section we will give some simulations of the unbiased covariance matrix estimate presented in Theorem 2.2 and 3.2.

4.1. Simulations of the regular normal distribution. In this section we assume that $\boldsymbol{x} \sim N_4\left(\mathbf{0}, \boldsymbol{\Sigma}_{(4)}^{(1)}\right)$, where

$$
\boldsymbol{\Sigma}_{(4)}^{(1)} = \begin{pmatrix} 5 & 2 & 0 & 0 \\ 2 & 5 & 1 & 0 \\ 0 & 1 & 5 & 3 \\ 0 & 0 & 3 & 5 \end{pmatrix}.
$$

In the simulation a sample of size $n = 20$ observations was randomly generated. Then the unbiased explicit estimates were calculated in each simulation. This was repeated 100000 times and the average values of the obtained estimate were calculated.

Based on the average explicit unbiased estimate is given by

$$
\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 4.99501 & 1.99590 & 0 & 0 \\ 1.99590 & 4.99238 & 0.99678 & 0 \\ 0 & 0.99678 & 5.00026 & 3.00265 \\ 0 & 0 & 3.00265 & 5.00368 \end{pmatrix}.
$$

In this simulation experiment the unbiased estimates seems to perform good.

4.2. Simulations of the estimators for a general linear model. In this section we assume the model $\boldsymbol{Y} = \boldsymbol{E} \sim N_{n,5}\left(\boldsymbol{BX}, \boldsymbol{\Sigma}_{(5)}^{(1)}, \boldsymbol{I}_n\right)$, where

$$\boldsymbol{\Sigma}_{(5)}^{(1)} = \begin{pmatrix} 4 & 1 & 0 & 0 & 0 \\ 1 & 3 & 2 & 0 & 0 \\ 0 & 2 & 5 & 3 & 0 \\ 0 & 0 & 3 & 5 & 3 \\ 0 & 0 & 0 & 3 & 5 \end{pmatrix}.$$

For each simulation the matrices $\boldsymbol{B}$ and $\boldsymbol{X}$ were randomly generated to avoid any effect on the estimation process.

In this simulation a sample of size $n = 80$ observations was randomly generated. Then the unbiased explicit estimates were calculated in each simulation. This was repeated 100000 times and the average values of the obtained estimate were calculated.

Based on the average explicit unbiased estimate is given by

$$\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 3.99865 & 0.99971 & 0 & 0 & 0 \\ 0.99971 & 3.00511 & 2.00246 & 0 & 0 \\ 0 & 2.00246 & 4.99898 & 2.99769 & 0 \\ 0 & 0 & 2.99769 & 4.99412 & 2.99504 \\ 0 & 0 & 0 & 2.99504 & 4.99352 \end{pmatrix}.$$

In this simulation experiment the unbiased estimates seems to perform good.

5. Conclusion

This paper presents unbiased and consistent estimators for a covariance matrix with a banded structure of order one. One can easily extend these results into a banded covariance matrix of any order. Similar results, as for the multivariate normal distribution, have also been shown for the general linear model. This new explicit estimator is more suitable to use in a real life situation since the property of unbiasedness is highly desired.

References

[1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Wiley-Interscience, Hoboken, NJ, 2003.

[2] S. A. Andersson, *Invariant normal models*, Ann. Statist. **3**(1) (1975), 132–154.

[3] J. P. Burg, D. G. Luenberger, and D. L. Wenger, *Estimation of structured covariance matrices*, Proceedings of the IEEE **70**(9) (1982), 963–974.

[4] M. Chakraborty, *An efficient algorithm for solving general periodic Toeplitz systems*, IEEE Trans. Signal Process. **46**(3) (1998), 784–787.

[5] L. P. B Christensen, *An EM-algorithm for band-Toeplitz covariance matrix estimation*, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Honolulu, Hawaii, USA, April 2007.

16

[6] D. R. Fuhrmann and T. A. Barton, *Estimation of block Toeplitz covariance matrices*, in: 24th IEEE Asilomar Conference on Signals, Systems and Computers. Pacific Grove, California, USA **2** (1990), pp. 779–783.

[7] J. M. Marin and T. Dhorne, *Linear Toeplitz covariance structure models with optimal estimators of variance components*, Linear Algebra Appl. **354**(1-3) (2002), 195–212.

[8] J. M. F. Moura and N. Balram, *Recursive structure of noncausal gauss-markov random fields*, IEEE Trans. Inform. Theory **38**(2) (1992), 334–354.

[9] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley-Interscience, New York, 1982.

[10] T. Nahtman, *Marginal permutation invariant covariance matrices with applications to linear models*, Linear Algebra Appl. **417**(1) (2006), 183–210.

[11] T. Nahtman and D. von Rosen, *Shift permutation invariance in linear random factor models.* Research Report Centre of Biostochastics, Swedish University of Agriculture science **2005:6** (2005).

[12] M. Ohlson, Z. Andrushchenko, and D. von Rosen, *Explicit estimators under m-dependence for a multivariate normal distribution*, Ann. Inst. Statist. Math. **63**(1) (2011), 29–42.

[13] I. Olkin, *Testing and estimation for structures which are circularly symmetric in blocks*, in: Multivariate Statistical Inference (D. G. Kabe and R. P. Gupta, eds.). North-Holland, Amsterdam, 1973, pp. 183–195.

[14] I. Olkin and S. J. Press, *Testing and estimation for a circular stationary model*, Ann. Math. Statist. **40**(4) (1969), 1358–1373.

[15] M. S. Srivastava and C. G. Khatri, *An Introduction to Multivariate Statistics*, North-Holland, New York – Oxford, 1979.

[16] D. F. Votaw, *Testing compound symmetry in a normal multivariate distribution*, Ann. Math. Statist. **19**(4) (1948), 447–473.

[17] S. S. Wilks, *Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution*, Ann. Math. Statist. **17**(3) (1946), 257–281.

[18] J. W. Woods, *Two-dimensional discrete markovian fields*, IEEE Trans. Inform. Theory **18**(2) (1972), 232–240.

Linköping university, Linköping, Sweden
E-mail address: `emika583@student.liu.se`
E-mail address: `martin.singull@liu.se`