Statistics and Applications {ISSN 2452-7395(online)}
Volume 18, No. 2, 2020 (New Series), pp 35–44

Upper and Lower Bounds of the Dispersion of a Mean Estimator in the Growth Curve Model

Shinpei Imori[1] and Dietrich von Rosen[2,3]
[1] *Graduate School of Advanced Science and Engineering*
Hiroshima University, Japan
[2] *Energy and Technology*
Swedish University of Agricultural Sciences
[3] *Department of Mathematics*
Linköping University, Sweden

Received: 26 March 2020; Revised: 15 April 2020; Accepted: 20 April 2020

Abstract

When more repeated measurements than independent observations are available the classical Growth Curve model cannot produce maximum likelihood estimators. In this article we are interested in the estimation of the mean parameters whereas the dispersion parameters are considered to be nuisance parameters. It is possible to produce an unbiased estimator of the mean parameters which is a function of the Moore-Penrose generalized inverse of a singular Wishart matrix. However, its dispersion seems very hard to derive. Therefore, upper and lower bounds of the dispersion are derived. Based on the bounds a general conclusion is that the proposed estimator will work better when the number of repeated measurements is much larger than the number of independent observations than when the number of repeated measurements and the number of independent observations are of similar size.

Key words: Growth curve model; High-dimensional setting; Moore-Penrose generalized inverse.

AMS2020 Subject Classifications: 62H12, 62H99

1. Introduction

High-dimensional statistics (more variables than independent observations) has been considered for many years and it is clear that many problems still are challenging to statistics. Kollo *et al.* (2011) published an article where the classical Growth Curve model (Potthoff & Roy, 1964) was treated and simulations indicated that the proposed ideas were reasonable. One overall conclusion was that problems with the Growth Curve model is more likely to occur when the number of variables is close to the number of independent observations but that the approach of Kollo *et al.* (2011) works when there are many more variables (repeated measurements) than independent observations. Unfortunately, the technical treatment was

Correponding Author: Dietrich von Rosen
Email: dietrich.von.rosen@slu.se

not completely correct since it was utilized that $\mathbf{A}(\mathbf{A}^\top\mathbf{V}\mathbf{A})^+\mathbf{A}^\top = \mathbf{V}^+$ for any non-singular matrix $\mathbf{A}$, $\mathbf{V}$ is positive semi-definite and $^+$ denotes the Moore-Penrose generalized inverse (for a definition of the inverse see *e.g.*, Kollo and von Rosen, 2005; Definition 1.1.5). This relation does however not hold unless $\mathbf{A}$ is an orthogonal matrix. In the paper by Kollo *et al.* (2011) there was a need to calculate moments of expressions involving the Moore-Penrose inverse of a singular Wishart matrix and this took place via the incorrect relation and some invariance arguments. In this article upper and lower bounds of the above mentioned moments will be derived which will support the overall conclusions in Kollo *et al.* (2011), although an exact expression for the dispersion matrix of the mean estimator seems very difficult to obtain.

Throughout the article bold upper cases will denote real matrices, $\mathbf{X} \sim N_{p,n}(\mathbf{0}, \mathbf{\Sigma}, \mathbf{I}_n)$ means that $\mathbf{X}$ of size $p \times n$ is matrix normally distributed with n independent columns which are multivariate normally distributed with mean equal to $\mathbf{0}$ and dispersion $\mathbf{\Sigma}$ which is supposed to be positive definite and $\mathbf{I}_n$ is the identity matrix of size $n \times n$. Note that the dispersion of $\mathbf{X}$ is given by $D[\mathbf{X}] = \mathbf{I}_n \otimes \mathbf{\Sigma}$, where $\otimes$ denotes the Kronecker product. Moreover, $\mathbf{V} \sim W_p(\mathbf{\Sigma}, n)$ denotes that $\mathbf{V}$ is Wishart distributed with n degrees of freedom, which holds if $\mathbf{V}$ can be factored as $\mathbf{V} = \mathbf{X}\mathbf{X}^\top$ (equality in distribution), where $\mathbf{X} \sim N_{p,n}(\mathbf{0}, \mathbf{\Sigma}, \mathbf{I}_n)$ and $^\top$ denotes the transpose. The rank of a matrix $\mathbf{A}$ is denoted $r(\mathbf{A})$.

2. Preparation

In this section three definitions and two useful lemmas are presented. Let $\mathbf{A} \geq 0$ ($\mathbf{A} > 0$) denote that $\mathbf{A}$ is positive semi-definite (positive definite) and let $\mathbf{A} \geq \mathbf{B}$ mean that $\mathbf{A} - \mathbf{B} \geq 0$, where both $\mathbf{A}$ and $\mathbf{B}$ are supposed to be positive semi-definite.

Definition 1:

(i) (Löwner ordering) Let $\mathbf{U}$ and $\mathbf{V}$ be positive semi-definite matrices. If for all vectors $\boldsymbol{\alpha}$ of proper size $\boldsymbol{\alpha}^\top\mathbf{U}\boldsymbol{\alpha} \leq \boldsymbol{\alpha}^\top\mathbf{V}\boldsymbol{\alpha}$ then $\mathbf{V} \geq \mathbf{U}$.

(ii) Let $\mathbf{U}$ and $\mathbf{V}$ be positive semi-definite matrices. If for all vectors $\boldsymbol{\alpha}$ of proper size $\boldsymbol{\alpha}^\top E[\mathbf{U}]\boldsymbol{\alpha} \leq \boldsymbol{\alpha}^\top E[\mathbf{V}]\boldsymbol{\alpha}$ then $E[\mathbf{V}] \geq E[\mathbf{U}]$.

(iii) If for all vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ of proper size

$$(\boldsymbol{\alpha} \otimes \boldsymbol{\beta})^\top D[\mathbf{X}](\boldsymbol{\alpha} \otimes \boldsymbol{\beta}) \leq (\boldsymbol{\alpha} \otimes \boldsymbol{\beta})^\top D[\mathbf{Y}](\boldsymbol{\alpha} \otimes \boldsymbol{\beta})$$

then it is written $D[\mathbf{X}] \preceq D[\mathbf{Y}]$, *i.e.*, $D[\boldsymbol{\beta}^\top\mathbf{X}\boldsymbol{\alpha}] \leq D[\boldsymbol{\beta}^\top\mathbf{Y}\boldsymbol{\alpha}]$.

The first lemma is presenting a known result of an explicit expression of a Moore-Penrose inverse of a singular Wishart matrix which can easily be verified via the four defining conditions of the Moore-Penrose inverse.

Lemma 1: Let $\mathbf{V} \sim W_p(\mathbf{\Sigma}, m)$, $p > m$. Then

$$\mathbf{V}^+ = \mathbf{U}(\mathbf{U}^\top\mathbf{U})^{-1}(\mathbf{U}^\top\mathbf{U})^{-1}\mathbf{U}^\top,$$

where $\mathbf{V} = \mathbf{U}\mathbf{U}^\top$ and $\mathbf{U} \sim N_{p,m}(\mathbf{0}, \mathbf{\Sigma}, \mathbf{I}_m)$.

Now a moment relation is presented which will be fundamental for the main results.

Lemma 2: Let $\mathbf{Q}$: $p \times q$, $q < p$, $\mathbf{P}$: $p \times p$ be of full rank and $\mathbf{V} \sim W_p(\mathbf{I}_p, n)$, $p < n$. Then

$$\begin{aligned} &E[(\mathbf{Q}^\top\mathbf{V}^{-1}\mathbf{Q})^{-1}\mathbf{Q}^\top\mathbf{V}^{-1}\mathbf{P}\mathbf{V}^{-1}\mathbf{Q}(\mathbf{Q}^\top\mathbf{V}^{-1}\mathbf{Q})^{-1} \\ &\quad = (\mathbf{Q}^\top\mathbf{Q})^{-1}\mathbf{Q}^\top\mathbf{P}\mathbf{Q}(\mathbf{Q}^\top\mathbf{Q})^{-1} + c_1\mathrm{tr}\{\mathbf{P}(\mathbf{I} - \mathbf{Q}(\mathbf{Q}^\top\mathbf{Q})^{-1}\mathbf{Q}^\top)\}(\mathbf{Q}^\top\mathbf{Q})^{-1}, \end{aligned}$$

where $c_1 = (n - (p - q) - 1)^{-1}$. **Proof:** Factor $\mathbf{Q}$ as $\mathbf{Q}^\top = \mathbf{H}(\mathbf{I}_q : \mathbf{0})\mathbf{\Gamma}$, where $\mathbf{H}$: $q \times q$ is a non-singular matrix and $\mathbf{\Gamma}$: $p \times p$ is an orthogonal matrix. Then,

$$\begin{aligned} &(\mathbf{Q}^\top\mathbf{V}^{-1}\mathbf{Q})^{-1}\mathbf{Q}^\top\mathbf{V}^{-1}\mathbf{P}\mathbf{V}^{-1}\mathbf{Q}(\mathbf{Q}^\top\mathbf{V}^{-1}\mathbf{Q})^{-1} \\ &\quad = (\mathbf{H}^\top)^{-1}((\mathbf{I}_q : \mathbf{0})\mathbf{\Gamma}\mathbf{V}^{-1}\mathbf{\Gamma}^\top(\mathbf{I}_q : \mathbf{0})^\top)^{-1}(\mathbf{I}_q : \mathbf{0})\mathbf{\Gamma}\mathbf{V}^{-1}\mathbf{P}\mathbf{V}^{-1}\mathbf{\Gamma}^\top(\mathbf{I}_q : \mathbf{0})^\top \\ &\qquad \times((\mathbf{I}_q : \mathbf{0})\mathbf{\Gamma}\mathbf{V}^{-1}\mathbf{\Gamma}^\top(\mathbf{I}_q : \mathbf{0})^\top)^{-1}\mathbf{H}^{-1}. \end{aligned} \tag{1}$$

Moreover, $\mathbf{\Gamma}\mathbf{V}^{-1}\mathbf{\Gamma}^\top = (\mathbf{\Gamma}\mathbf{V}\mathbf{\Gamma}^\top)^{-1}$ follows the same distribution as $\mathbf{V}^{-1}$. Thus, the right hand side of (1) follows the same distribution as

$$(\mathbf{H}^\top)^{-1}(\mathbf{V}^{11})^{-1}(\mathbf{V}^{11} : \mathbf{V}^{12})\mathbf{\Gamma}\mathbf{P}\mathbf{\Gamma}^\top(\mathbf{V}^{11} : \mathbf{V}^{12})^\top(\mathbf{V}^{11})^{-1}\mathbf{H}^{-1} \tag{2}$$

where $\mathbf{V}^{11}$ and $\mathbf{V}^{12}$ are defined via

$$\mathbf{V}^{-1} = \begin{pmatrix} \mathbf{V}^{11} & \mathbf{V}^{12} \\ \mathbf{V}^{21} & \mathbf{V}^{22} \end{pmatrix}, \qquad \begin{matrix} q \times q & q \times (p-q) \\ (p-q) \times q & (p-q) \times (p-q) \end{matrix}$$

and similarly $\mathbf{V}_{12}$ and $\mathbf{V}_{22}$ are defined through

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{pmatrix}, \qquad \begin{matrix} q \times q & q \times (p-q) \\ (p-q) \times q & (p-q) \times (p-q) \end{matrix}.$$

The submatrices satisfy (see *e.g.* Kollo and von Rosen, 2005; Proposition 1.3.4 (i)) $(\mathbf{V}^{11})^{-1}\mathbf{V}^{12} = -\mathbf{V}_{12}\mathbf{V}_{22}^{-1}$. Let $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$ be defined through $\mathbf{\Gamma}^\top = (\mathbf{\Gamma}_1^\top : \mathbf{\Gamma}_2^\top)$: $(p \times q : p \times (p-q))$ and note that $\mathbf{\Gamma}_1\mathbf{\Gamma}_1^\top = \mathbf{I}_q$. Then (2) equals

$$\begin{aligned} &(\mathbf{H}^\top)^{-1}(\mathbf{\Gamma}_1 - \mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{\Gamma}_2)\mathbf{P}(\mathbf{\Gamma}_1^\top - \mathbf{\Gamma}_2^\top\mathbf{V}_{22}^{-1}\mathbf{V}_{12})\mathbf{H}^{-1} \\ &\quad = (\mathbf{H}^\top)^{-1}(\mathbf{\Gamma}_1\mathbf{P}\mathbf{\Gamma}_1^\top - \mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{\Gamma}_2\mathbf{P}\mathbf{\Gamma}_1^\top - \mathbf{\Gamma}_1\mathbf{P}\mathbf{\Gamma}_2^\top\mathbf{V}_{22}^{-1}\mathbf{V}_{21} \\ &\qquad + \mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{\Gamma}_2\mathbf{P}\mathbf{\Gamma}_2^\top\mathbf{V}_{22}^{-1}\mathbf{V}_{21})\mathbf{H}^{-1}. \end{aligned} \tag{3}$$

It will be utilized that $\mathbf{V}_{12}\mathbf{V}_{22}^{-1/2} \sim N_{q,p-q}(\mathbf{0}, \mathbf{I}_q, \mathbf{I}_{p-q})$ which is independent of $\mathbf{V}_{22}$ (see *e.g.*, Kollo and von Rosen, 2005; Theorem 2.4.12). The expectation of the expression in the right hand side of (3) is to be derived. Since $E[\mathbf{V}_{12}\mathbf{V}_{22}^{-1}] = \mathbf{0}$ it follows from (3) that the next expression should be calculated:

$$(\mathbf{H}^\top)^{-1}\mathbf{\Gamma}_1\mathbf{P}\mathbf{\Gamma}_1\mathbf{H}^{-1} + (\mathbf{H}^\top)^{-1}E[\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{\Gamma}_2\mathbf{P}\mathbf{\Gamma}_2^\top\mathbf{V}_{22}^{-1}\mathbf{V}_{21}]\mathbf{H}^{-1}. \tag{4}$$

Moreover, applying an expectation result for quadratic forms in normally distributed variables (*e.g.*, see Kollo and von Rosen, 2005, Theorem 2.2.9 (i)) implies that the expectation in (4) equals

$$E[\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{\Gamma}_2\mathbf{P}\mathbf{\Gamma}_2^\top\mathbf{V}_{22}^{-1}\mathbf{V}_{21}] = E[\mathrm{tr}\{\mathbf{V}_{22}^{-1}\mathbf{\Gamma}_2\mathbf{P}\mathbf{\Gamma}_2^\top\}]\mathbf{I}_\mathrm{q} \tag{5}$$

which since $\mathbf{V}_{22} \sim W_{p-q}(\mathbf{I}, n)$ (see *e.g.*, Kollo and von Rosen, 2005, Theorem 2.4.14 (iii))

$$E[\mathbf{V}_{22}^{-1}] = c_1 \mathbf{I}_{p-q}, \qquad c_1 = \tfrac{1}{n-(p-q)-1}$$

and the right hand side of (5) is identical to $c_1 E[\mathrm{tr}\{\mathbf{\Gamma}_2 \mathbf{P} \mathbf{\Gamma}_2^\top\}]\mathbf{I}_\mathrm{q}$. In order to arrive to the statements of the theorem $\mathbf{\Gamma}_2^\top \mathbf{\Gamma}_2$, $(\mathbf{H}\mathbf{H}^\top)^{-1}$ and $\mathbf{\Gamma}_1^\top \mathbf{H}^{-1}$ have to be expressed in the original matrices. From the definition of $\mathbf{\Gamma}$ and $\mathbf{H}$ it follows that

$$\mathbf{Q}^\top = \mathbf{H}\mathbf{\Gamma}_1, \quad (\mathbf{H}\mathbf{H}^\top)^{-1} = (\mathbf{Q}^\top\mathbf{Q})^{-1}, \quad \mathbf{\Gamma}_1^\top \mathbf{H}^{-1} = \mathbf{Q}(\mathbf{Q}^\top\mathbf{Q})^{-1}$$

and

$$\mathbf{\Gamma}_1^\top \mathbf{\Gamma}_1 = \mathbf{Q}(\mathbf{Q}^\top\mathbf{Q})^{-1}\mathbf{Q}^\top, \quad \mathbf{\Gamma}_2^\top \mathbf{\Gamma}_2 = \mathbf{I} - \mathbf{\Gamma}_1^\top \mathbf{\Gamma}_1 = \mathbf{I} - \mathbf{Q}(\mathbf{Q}^\top\mathbf{Q})^{-1}\mathbf{Q}^\top.$$

These relations establish the lemma. □ Throughout the article let $\lambda_1(\mathbf{A}) \geq \lambda_2(\mathbf{A}) \geq \cdots \geq \lambda_n(\mathbf{A})$ be the ordered eigen values of a symmetric matrix $\mathbf{A}$: $n \times n$.

3. The Growth Curve Model When $p > n$

The classical Growth Curve model (Potthoff and Roy, 1964) has been applied in many areas and is a natural extension of the MANOVA model. Therefore the model is also called GMANOVA model. The model and generalizations of the model together with an almost complete list of references can be found in von Rosen (2018).

Definitiion 2: Let $\mathbf{X}$: $p \times n$, $p \geq n - r(\mathbf{C})$, $\mathbf{A}$: $p \times q$, $q \leq p$, $\mathbf{B}$: $q \times k$, $\mathbf{C}$: $k \times n$, and $\mathbf{\Sigma} > 0$: $p \times p$. Then

$$\mathbf{X} = \mathbf{A}\mathbf{B}\mathbf{C} + \mathbf{E}$$

defines the Growth Curve model, where $\mathbf{E} \sim N_{p,n}(\mathbf{0}, \mathbf{\Sigma}, \mathbf{I})$, $\mathbf{A}$ and $\mathbf{C}$ are known matrices, and $\mathbf{B}$ and $\mathbf{\Sigma}$ are unknown parameter matrices.

Since $p \geq n - r(\mathbf{C})$ we assume a high-dimensional setting. The main purpose of this article is to discuss a specific estimator of $\mathbf{B}$. Note that the size of $\mathbf{B}$, *i.e.*, $q \times k$, does not depend on n and p, and that $\mathbf{\Sigma}$ is thought of being a nuisance parameter. For some details of how to treat the Growth Curve model in a high-dimensional setting see Kollo *et al.* (2011). Alternatively, we can suppose that $\mathbf{\Sigma}$ is known and then from linear models theory it follows that under the assumption $r(\mathbf{A}) = q$, $r(\mathbf{C}) = k$ which will be supposed to hold throughout the article, an estimator of $\mathbf{B}$ equals

$$\widetilde{\mathbf{B}} = (\mathbf{A}^\top \mathbf{\Sigma}^{-1} \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{\Sigma}^{-1} \mathbf{X} \mathbf{C}^\top (\mathbf{C}\mathbf{C}^\top)^{-1}.$$

However, $\mathbf{\Sigma}^{-1}$ has to be estimated. One can use the sums of squares matrix $\mathbf{S} = \mathbf{X}(\mathbf{I} - \mathbf{C}^\top(\mathbf{C}\mathbf{C}^\top)^{-1}\mathbf{C})\mathbf{X}^\top$ as an estimator of $n\mathbf{\Sigma}$ but if $p > n - r(\mathbf{C})$ the inverse $\mathbf{S}^{-1}$ does not exist and cannot be used to estimate $\mathbf{\Sigma}^{-1}$. Therefore, instead of $\mathbf{S}^{-1}$ the Moore-Penrose inverse $\mathbf{S}^{+}$ can be used and then the same estimator as in Kollo *et al.* (2011) is obtained:

$$\widehat{\mathbf{B}} = (\mathbf{A}^\top \mathbf{S}^{+} \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{S}^{+} \mathbf{X} \mathbf{C}^\top (\mathbf{C}\mathbf{C}^\top)^{-1}, \tag{6}$$

where it has to be assumed that the column vector space relation $\mathcal{C}(\mathbf{A}) \cap \mathcal{C}(\mathbf{S}) = \{\mathbf{0}\}$ is satisfied which implies that $(\mathbf{A}^\top\mathbf{S}^+\mathbf{A})^{-1}$ holds.

4. $E[\widehat{\mathbf{B}}]$ and Bounds For $D[\widehat{\mathbf{B}}]$

In order to derive the expectation and bounds for the dispersion for $\widehat{\mathbf{B}}$ in (6) it will be utilized that $\mathbf{X}\mathbf{C}^\top$ and $\mathbf{S}$ are independently distributed.

Theorem 1: Let $\widehat{\mathbf{B}}$ be defined in (6). Then $E[\widehat{\mathbf{B}}] = \mathbf{B}$.

Proof:

$$E[\widehat{\mathbf{B}}] = E[(\mathbf{A}^\top\mathbf{S}^+\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{S}^+]E[\mathbf{X}\mathbf{C}^\top(\mathbf{C}\mathbf{C}^\top)^{-1}] = E[(\mathbf{A}^\top\mathbf{S}^+\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{S}^+\mathbf{A}\mathbf{B}] = \mathbf{B}.$$

□

Turning to a discussion of the dispersion matrix for $\widehat{\mathbf{B}}$ it follows that the dispersion matrix for $\widehat{\mathbf{B}}$ in (6) can be presented as

$$\begin{aligned} D[\widehat{\mathbf{B}}] &= E[\mathrm{vec}(\widehat{\mathbf{B}}-\mathbf{B})\mathrm{vec}^\top(\widehat{\mathbf{B}}-\mathbf{B})] \\ &= E[((\mathbf{C}\mathbf{C}^\top)^{-1}\mathbf{C}\otimes(\mathbf{A}^\top\mathbf{S}^+\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{S}^+)D[\mathbf{X}](\mathbf{C}^\top(\mathbf{C}\mathbf{C}^\top)^{-1}\otimes\mathbf{S}^+\mathbf{A}(\mathbf{A}^\top\mathbf{S}^+\mathbf{A})^{-1})] \\ &= (\mathbf{C}\mathbf{C}^\top)^{-1}\otimes E[(\mathbf{A}^\top\mathbf{S}^+\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{S}^+\boldsymbol{\Sigma}\mathbf{S}^+\mathbf{A}(\mathbf{A}^\top\mathbf{S}^+\mathbf{A})^{-1}]. \end{aligned} \tag{7}$$

From Lemma 1 it follows that the expectation in (7) is complicated to express. We will show some calculations but the aim will be to find upper and lower bounds for the expectation similarly to the approach for obtaining bounds for the expectation and dispersion of the Moore-Penrose inverse of a singular Wishart matrix (see Imori and von Rosen, 2020).

When deriving the bounds a number of transformations will take place: $\mathbf{S} = \mathbf{U}\mathbf{U}^\top$, where $\mathbf{U} \sim N_{p,n-r(\mathbf{C})}(\mathbf{0},\boldsymbol{\Sigma},\mathbf{I}_{n-r(\mathbf{C})})$, $\mathbf{Y} = \boldsymbol{\Sigma}^{-1/2}\mathbf{U}$, where $\boldsymbol{\Sigma}^{-1/2}$ is a symmetric square root; $\mathbf{Y}^\top = \mathbf{T}\mathbf{L}$ where $\mathbf{T}$: $(n-r(\mathbf{C}))\times(n-r(\mathbf{C}))$ is a lower triangular matrix and $\mathbf{L}$: $(n-r(\mathbf{C}))\times p$ is a semi-orthogonal matrix, *i.e.* $\mathbf{L}\mathbf{L}^\top = \mathbf{I}_{n-r(\mathbf{C})}$; $\mathbf{V} = \mathbf{T}^\top\mathbf{T}$. Firstly it is noted that the expectation in (7) can be expressed as (see Lemma 1)

$$\begin{aligned} &E[(\mathbf{A}^\top\mathbf{S}^+\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{S}^+\boldsymbol{\Sigma}\mathbf{S}^+\mathbf{A}(\mathbf{A}^\top\mathbf{S}^+\mathbf{A})^{-1}] \\ &\quad = E[(\mathbf{A}^\top\mathbf{U}(\mathbf{U}^\top\mathbf{U})^{-1}(\mathbf{U}^\top\mathbf{U})^{-1}\mathbf{U}^\top\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{U}(\mathbf{U}^\top\mathbf{U})^{-1}(\mathbf{U}^\top\mathbf{U})^{-1}\mathbf{U}^\top\boldsymbol{\Sigma} \\ &\qquad \times\mathbf{U}(\mathbf{U}^\top\mathbf{U})^{-1}(\mathbf{U}^\top\mathbf{U})^{-1}\mathbf{U}^\top\mathbf{A}(\mathbf{A}^\top\mathbf{U}(\mathbf{U}^\top\mathbf{U})^{-1}(\mathbf{U}^\top\mathbf{U})^{-1}\mathbf{U}^\top\mathbf{A})^{-1}]. \end{aligned}$$

Now the first transformation is applied to this relation and it yields

$$\begin{aligned} &E[(\mathbf{A}^\top\boldsymbol{\Sigma}^{1/2}\mathbf{Y}(\mathbf{Y}^\top\boldsymbol{\Sigma}\mathbf{Y})^{-1}(\mathbf{Y}^\top\boldsymbol{\Sigma}\mathbf{Y})^{-1}\mathbf{Y}^\top\boldsymbol{\Sigma}^{1/2}\mathbf{A})^{-1} \\ &\quad\times\mathbf{A}^\top\boldsymbol{\Sigma}^{1/2}\mathbf{Y}(\mathbf{Y}^\top\boldsymbol{\Sigma}\mathbf{Y})^{-1}(\mathbf{Y}^\top\boldsymbol{\Sigma}\mathbf{Y})^{-1}\mathbf{Y}^\top\boldsymbol{\Sigma} \\ &\quad\times\boldsymbol{\Sigma}\mathbf{Y}(\mathbf{Y}^\top\boldsymbol{\Sigma}\mathbf{Y})^{-1}(\mathbf{Y}^\top\boldsymbol{\Sigma}\mathbf{Y})^{-1}\mathbf{Y}^\top\boldsymbol{\Sigma}^{1/2}\mathbf{A} \\ &\quad\times(\mathbf{A}^\top\boldsymbol{\Sigma}^{1/2}\mathbf{Y}(\mathbf{Y}^\top\boldsymbol{\Sigma}\mathbf{Y})^{-1}(\mathbf{Y}^\top\boldsymbol{\Sigma}\mathbf{Y})^{-1}\mathbf{Y}^\top\boldsymbol{\Sigma}^{1/2}\mathbf{A})^{-1}]. \end{aligned}$$

Moreover, the second transformation implies that one should consider

$$\begin{aligned} &E[(\mathbf{A}^\top\boldsymbol{\Sigma}^{1/2}\mathbf{L}^\top(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{T}^{-1}(\mathbf{T}^\top)^{-1}(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{L}\boldsymbol{\Sigma}^{1/2}\mathbf{A})^{-1} \\ &\quad\times\mathbf{A}^\top\boldsymbol{\Sigma}^{1/2}\mathbf{L}^\top(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{T}^{-1}(\mathbf{T}^\top)^{-1}(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{L}\boldsymbol{\Sigma} \\ &\quad\times\boldsymbol{\Sigma}\mathbf{L}^\top(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{T}^{-1}(\mathbf{T}^\top)^{-1}(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{L}\boldsymbol{\Sigma}^{1/2}\mathbf{A} \\ &\quad\times(\mathbf{A}^\top\boldsymbol{\Sigma}^{1/2}\mathbf{L}^\top(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{T}^{-1}(\mathbf{T}^\top)^{-1}(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{L}\boldsymbol{\Sigma}^{1/2}\mathbf{A})^{-1}] \end{aligned}$$

and then the third transformation implies an expression which will be studied in detail:

$$\begin{aligned}
&E[(\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{L}^\top(\mathbf{L}\mathbf{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{V}^{-1}(\mathbf{L}\mathbf{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{L}\mathbf{\Sigma}^{1/2}\mathbf{A})^{-1} \\
&\quad\times\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{L}^\top(\mathbf{L}\mathbf{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{V}^{-1}(\mathbf{L}\mathbf{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{L}\mathbf{\Sigma} \\
&\quad\times\mathbf{\Sigma}\mathbf{L}^\top(\mathbf{L}\mathbf{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{V}^{-1}(\mathbf{L}\mathbf{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{L}\mathbf{\Sigma}^{1/2}\mathbf{A} \\
&\quad\times(\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{L}^\top(\mathbf{L}\mathbf{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{V}^{-1}(\mathbf{L}\mathbf{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{L}\mathbf{\Sigma}^{1/2}\mathbf{A})^{-1}], \qquad (8)
\end{aligned}$$

where it can be shown that $\mathbf{V} \sim W_{n-r(\mathbf{C})}(\mathbf{I}_{n-r(\mathbf{C})}, p)$ which for example follows from the derivation of the Wishart density in Srivastava and Khatri (1979; Theorem 3.2.1) or Imori and von Rosen (2020; Section 3.1). Put

$$\begin{aligned}
\mathbf{Q} &= (\mathbf{L}\mathbf{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{L}\mathbf{\Sigma}^{1/2}\mathbf{A}, \qquad (9)\\
\mathbf{P} &= (\mathbf{L}\mathbf{\Sigma}\mathbf{L}^\top)^{-1}\mathbf{L}\mathbf{\Sigma}\mathbf{\Sigma}\mathbf{L}^\top(\mathbf{L}\mathbf{\Sigma}\mathbf{L}^\top)^{-1}. \qquad (10)
\end{aligned}$$

Then (8) is identical to

$$E[(\mathbf{Q}^\top\mathbf{V}^{-1}\mathbf{Q})^{-1}\mathbf{Q}^\top\mathbf{V}^{-1}\mathbf{P}\mathbf{V}^{-1}\mathbf{Q}(\mathbf{Q}^\top\mathbf{V}^{-1}\mathbf{Q})^{-1}]$$

which, since $\mathbf{V} \sim W_{n-r(\mathbf{C})}(\mathbf{I}_{n-r(\mathbf{C})}, p)$, according to Lemma 2

$$E[(\mathbf{Q}^\top\mathbf{Q})^{-1}\mathbf{Q}^\top\mathbf{P}\mathbf{Q}(\mathbf{Q}^\top\mathbf{Q})^{-1}] + c_1E[\mathrm{tr}\{\mathbf{P}(\mathbf{I}-\mathbf{Q}(\mathbf{Q}^\top\mathbf{Q})^{-1}\mathbf{Q}^\top)\}(\mathbf{Q}^\top\mathbf{Q})^{-1}], \qquad (11)$$

where $c_1^{-1} = p - (n - r(\mathbf{C}) - q) - 1$ and the expectation in (11) is taken over the semi-orthogonal matrix $\mathbf{L}$. Note that it has to be assumed that $c_1 > 0$, *i.e.*, $p > n - r(\mathbf{C}) - q + 1$ but later we need that $p \geq n - r(\mathbf{C})$) . However, it is difficult to perform the integration in (11) and therefore we first focus on finding upper and lower bounds of

$$(\mathbf{Q}^\top\mathbf{Q})^{-1}\mathbf{Q}^\top\mathbf{P}\mathbf{Q}(\mathbf{Q}^\top\mathbf{Q})^{-1}, \quad \mathrm{tr}\{\mathbf{P}(\mathbf{I}-\mathbf{Q}(\mathbf{Q}^\top\mathbf{Q})^{-1}\mathbf{Q}^\top)\}$$

which are either functionally independent of $\mathbf{L}$ or are so simplified that only $E[(\mathbf{Q}^\top\mathbf{Q})^{-1}]$ has to be derived.

Lemma 3: Let $\mathbf{P}$ be given by (10). Then

$$\lambda_p(\mathbf{\Sigma})\lambda_p(\mathbf{\Sigma}^{-1})\mathbf{I}_{n-r(\mathbf{C})} \leq \mathbf{P} \leq \lambda_1(\mathbf{\Sigma})\lambda_1(\mathbf{\Sigma}^{-1})\mathbf{I}_{n-r(\mathbf{C})}.$$

Proof: The proof is based on a spectral decomposition of $\mathbf{\Sigma}$ which yields $\lambda_p(\mathbf{\Sigma})\mathbf{I}_p \leq \mathbf{\Sigma} \leq \lambda_1(\mathbf{\Sigma})\mathbf{I}_p$. Note that $\lambda_p(\mathbf{\Sigma})\mathbf{\Sigma} \leq \mathbf{\Sigma}\mathbf{\Sigma} \leq \lambda_1(\mathbf{\Sigma})\mathbf{\Sigma}$ and therefore $\lambda_1(\mathbf{L}\mathbf{\Sigma}\mathbf{L}^\top) \leq \lambda_1(\mathbf{\Sigma})$, $\lambda_n(\mathbf{L}\mathbf{\Sigma}\mathbf{L}^\top) \geq \lambda_p(\mathbf{\Sigma})$ which jointly establish the lemma. $\square$ Applying Lemma 3 yields that upper and lower bounds for (11) are given by

$$\begin{aligned}
&\lambda_p(\mathbf{\Sigma})\lambda_p(\mathbf{\Sigma}^{-1})(1 + c_1(n - r(\mathbf{C}) - q))E[(\mathbf{Q}\mathbf{Q}^\top)^{-1}] \\
&\quad\leq E[(\mathbf{Q}^\top\mathbf{Q})^{-1}\mathbf{Q}^\top\mathbf{P}\mathbf{Q}(\mathbf{Q}^\top\mathbf{Q})^{-1}] + c_1E[\mathrm{tr}\{\mathbf{P}(\mathbf{I}-\mathbf{Q}(\mathbf{Q}^\top\mathbf{Q})^{-1}\mathbf{Q}^\top)\}(\mathbf{Q}^\top\mathbf{Q})^{-1}] \\
&\quad\leq \lambda_1(\mathbf{\Sigma})\lambda_1(\mathbf{\Sigma}^{-1})(1 + c_1(n - r(\mathbf{C}) - q))E[(\mathbf{Q}\mathbf{Q}^\top)^{-1}]. \qquad (12)
\end{aligned}$$

Note that $1 + c_1(n - r(\mathbf{C}) - q) = (p-1)/(n - r(\mathbf{C}) - q)$. Moreover, (12) implies that we now need to find bounds for $E[(\mathbf{Q}\mathbf{Q}^\top)^{-1}]$.

Lemma 4: Let $\mathbf{Q}$ be defined in (9). Then

$$\begin{aligned}\lambda_p(\mathbf{\Sigma\Sigma})(\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{L}^\top\mathbf{L}\mathbf{\Sigma}^{1/2}\mathbf{A})^{-1} &\leq (\mathbf{Q}^\top\mathbf{Q})^{-1}\\ &\leq \lambda_1(\mathbf{\Sigma\Sigma})(\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{L}^\top\mathbf{L}\mathbf{\Sigma}^{1/2}\mathbf{A})^{-1}.\end{aligned}$$

Proof: It is enough to show upper and lower limits for $(\mathbf{L\Sigma L}^\top)^{-1}(\mathbf{L\Sigma L}^\top)^{-1}$ which should be independent of $\mathbf{L}$ and proportional to $\mathbf{I}_{n-r(\mathbf{C})}$. By pre- and post-multiplying by $\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{L}^\top$ and then taking the inverse establish the lemma. Now

$$\lambda_n((\mathbf{L\Sigma L}^\top)^{-1}(\mathbf{L\Sigma L}^\top)^{-1}) = (\lambda_1(\mathbf{L\Sigma L}^\top))^{-2} \geq (\lambda_1(\mathbf{\Sigma\Sigma}))^{-1}$$

and

$$\lambda_1((\mathbf{L\Sigma L}^\top)^{-1}(\mathbf{L\Sigma L}^\top)^{-1}) = (\lambda_n(\mathbf{L\Sigma L}^\top))^{-2} \leq (\lambda_p(\mathbf{\Sigma\Sigma}))^{-1}$$

which yield the inequalities of the lemma. □ From Lemma 4 it follows that we need to calculate $E[(\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{L}^\top\mathbf{L}\mathbf{\Sigma}^{1/2}\mathbf{A})^{-1}]$, where $\mathbf{LL}^\top = \mathbf{I}_{n-r(\mathbf{C})}$. The result is stated in the next lemma.

Lemma 5: Let all matrices be as in Lemma 4. Then, if $p \geq n - r(\mathbf{C}) > q - 1$,

$$E[(\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{L}^\top\mathbf{L}\mathbf{\Sigma}^{1/2}\mathbf{A})^{-1}] = \frac{p-q-1}{n-r(\mathbf{C})-q-1}(\mathbf{A}^\top\mathbf{\Sigma}\mathbf{A})^{-1}.$$

Proof: The same transformations as when deriving (8) will now be applied. Let $\mathbf{Y} \sim N_{p,n-r(\mathbf{C})}(\mathbf{0}, \mathbf{I}_p, \mathbf{I}_{n-r(\mathbf{C})})$, $p \geq n - r(\mathbf{C})$. Then $\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{YY}^\top\mathbf{\Sigma}^{1/2}\mathbf{A} \sim W_q(\mathbf{A}^\top\mathbf{\Sigma}\mathbf{A}, n - r(\mathbf{C}))$ and (expectation of an inverse Wishart matrix is applied; *e.g.*, see Kollo and von Rosen, 2005, Theorem 2.4.14 (iii))

$$E[(\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{YY}^\top\mathbf{\Sigma}^{1/2}\mathbf{A})^{-1}] = \frac{1}{n-r(\mathbf{C})-q-1}(\mathbf{A}^\top\mathbf{\Sigma}\mathbf{A})^{-1}. \tag{13}$$

Next the variable substitution $\mathbf{Y}^\top = \mathbf{TL}$ is made, where $\mathbf{T}$: $(n-r(\mathbf{C}))\times(n-r(\mathbf{C}))$, is lower triangular with positive diagonal elements and $\mathbf{L}$ is semi-orthogonal, *i.e.*, $\mathbf{LL}^\top = \mathbf{I}_{n-r(\mathbf{C})}$. The matrices $\mathbf{T}$ and $\mathbf{L}$ are independently distributed. Moreover, $\mathbf{V} = \mathbf{T}^\top\mathbf{T} \sim W_{n-r(\mathbf{C})}(\mathbf{I}_{n-r(\mathbf{C})}, p)$ and given $\mathbf{L}$

$$\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{L}^\top\mathbf{VL}\mathbf{\Sigma}^{1/2}\mathbf{A} \sim W_q(\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{L}^\top\mathbf{L}\mathbf{\Sigma}^{1/2}\mathbf{A}, p).$$

Thus,

$$\begin{aligned}E[(\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{YY}^\top\mathbf{\Sigma}^{1/2}\mathbf{A})^{-1}] &= E[(\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{L}^\top\mathbf{T}^\top\mathbf{TL}\mathbf{\Sigma}^{1/2}\mathbf{A})^{-1}]\\ = E[(\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{L}^\top\mathbf{VL}\mathbf{\Sigma}^{1/2}\mathbf{A})^{-1}] &= \frac{1}{p-q-1}E[(\mathbf{A}^\top\mathbf{\Sigma}^{1/2}\mathbf{L}^\top\mathbf{L}\mathbf{\Sigma}^{1/2}\mathbf{A})^{-1}]\end{aligned}$$

and combining this result with (13) establishes the lemma □ From Lemma 3 and Lemma 5 it follows that

$$\begin{aligned}&\lambda_p(\mathbf{\Sigma})\lambda_p(\mathbf{\Sigma}^{-1})\lambda_p(\mathbf{\Sigma\Sigma})(1 + c_1(n-r(\mathbf{C})-q)\tfrac{p-q-1}{n-r(\mathbf{C})-q-1})(\mathbf{A}^\top\mathbf{\Sigma}\mathbf{A})^{-1}\\ &\quad\leq E[(\mathbf{A}^\top\mathbf{S}^+\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{S}^+\mathbf{\Sigma}\mathbf{S}^+\mathbf{A}(\mathbf{A}^\top\mathbf{S}^+\mathbf{A})^{-1}]\\ &\quad\leq \lambda_1(\mathbf{\Sigma})\lambda_1(\mathbf{\Sigma}^{-1})\lambda_1(\mathbf{\Sigma\Sigma})(1 + c_1(n-r(\mathbf{C})-q)\tfrac{p-q-1}{n-r(\mathbf{C})-q-1})(\mathbf{A}^\top\mathbf{\Sigma}\mathbf{A})^{-1}.\end{aligned} \tag{14}$$

Now all preparations are finished and the main result can immediately be presented:

Theorem 2: Let $\widehat{\mathbf{B}}$ be defined in (6) and assume $p \geq n - r(\mathbf{C}) > q - 1$. Then, ($\succeq$ was introduced in Definition 1 (iii))

(i) $D[\widehat{\mathbf{B}}] \succeq (\mathbf{C}\mathbf{C}^\top)^{-1} \otimes \lambda_p(\mathbf{\Sigma})^3 \lambda_p(\mathbf{\Sigma}^{-1}) \frac{p-1}{p-(n-r(\mathbf{C})-q)-1} \frac{p-q-1}{n-r(\mathbf{C})-q-1} (\mathbf{A}^\top \mathbf{\Sigma} \mathbf{A})^{-1}$;

(ii) $D[\widehat{\mathbf{B}}] \preceq (\mathbf{C}\mathbf{C}^\top)^{-1} \otimes \lambda_1(\mathbf{\Sigma})^3 \lambda_1(\mathbf{\Sigma}^{-1}) \frac{p-1}{p-(n-r(\mathbf{C})-q)-1} \frac{p-q-1}{n-r(\mathbf{C})-q-1} (\mathbf{A}^\top \mathbf{\Sigma} \mathbf{A})^{-1}$.

Remark 1: If $p-1$ is close to $n-r(\mathbf{C})-q$ or $n-r(\mathbf{C})$ is close to $q-1$ the dispersion for $\widehat{\mathbf{B}}$ becomes large because the lower bound becomes large. In this case an alternative estimator for $\mathbf{B}$ should be used, *e.g.*, the "unweighted" estimator $(\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{X}\mathbf{C}^\top(\mathbf{C}\mathbf{C}^\top)^{-1}$.

If all eigen values of $\mathbf{\Sigma}$ are equal, *e.g.*, $\mathbf{\Sigma} = \mathbf{I}$, the lower and upper bound of Theorem 2 are equal, *i.e.*,

$$D[\widehat{\mathbf{B}}] = \lambda_1(\mathbf{\Sigma}) \frac{p-1}{p-(n-r(\mathbf{C})-q)-1} \frac{p-q-1}{n-r(\mathbf{C})-q-1} (\mathbf{C}\mathbf{C}^\top)^{-1} \otimes (\mathbf{A}^\top\mathbf{A})^{-1}$$

which however is larger than the variance for the unweighted estimator, as it should be according to least squares theory.

5. Simulation Study

In this section a small simulation study is conducted to illustrate Theorem 1. In Remark 1 it was noted that when p is close to n both the upper and lower bounds, for given $(\mathbf{C}\mathbf{C}^\top)^{-1} \otimes (\mathbf{A}^\top\mathbf{\Sigma}\mathbf{A})^{-1}$, depend on

$$\lambda_\bullet(\mathbf{\Sigma})^3 \lambda_\bullet(\mathbf{\Sigma}^{-1}) \frac{p-1}{p-(n-r(\mathbf{C})-q)-1} \frac{p-q-1}{n-r(\mathbf{C})-q-1}, \qquad p \geq n - r(\mathbf{C}),$$

where $\lambda_\bullet(\mathbf{\Sigma})$ denotes either $\lambda_1(\mathbf{\Sigma})$ or $\lambda_p(\mathbf{\Sigma})$ and the same holds for $\lambda_\bullet(\mathbf{\Sigma}^{-1})$. If $p = n - r(\mathbf{C})$ this expression reduces to

$$\lambda_\bullet(\mathbf{\Sigma})^3 \lambda_\bullet(\mathbf{\Sigma}^{-1}) \frac{p-1}{q-1}. \tag{15}$$

Thus, if the largest and smallest eigenvalues of $\mathbf{\Sigma}$ are stable with respect to p (15) increases linearly with p but at the same time $(\mathbf{A}^\top\mathbf{\Sigma}\mathbf{A})^{-1}$ becomes "smaller". Note also that $(\mathbf{C}\mathbf{C}^\top)^{-1}$ becomes "smaller" when n increases.

Instead of studying Theorem 1 we will study (14) since $(\mathbf{C}\mathbf{C}^\top)^{-1}$ is of no interest. In the simulations the following matrices are used: $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$, where

$$\mathbf{a}_1 = \mathbf{1}_p, \qquad \mathbf{a}_2^\top = 0.7 \cdot (1, 2, \ldots, p), \qquad \mathbf{a}_3^\top = 0.01 \cdot (1, 4, \ldots, p^2)$$

and

$$\mathbf{C} = \begin{pmatrix} \mathbf{1}_{20}^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{20}^\top \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 3 & 7 \\ 2 & 2 \end{pmatrix}.$$

Table 1: The table summarizes the conducted simulation study: $(\mathbf{EST})_{ii}$, $i \in \{1,2,3\}$ is the ith diagonal element of EST defined in (16), $(\mathbf{LB})_{ii}$ and $(\mathbf{UB})_{ii}$, $i \in \{1,2,3\}$, are the ith diagonal element of the lower and upper bounds respectively defined in (18) and (19). Moreover, p is the number of repeated measurements and the data has been generated according to the description in Section 5. In particular $n = 40$, $q = 3$ and $r(\mathbf{C}) = 2$.

p	$(\mathbf{LB})_{11}$	$(\mathbf{EST})_{11}$	$(\mathbf{UB})_{11}$	$(\mathbf{LB})_{22}$	$(\mathbf{EST})_{22}$	$(\mathbf{UB})_{22}$	$(\mathbf{LB})_{33}$	$(\mathbf{EST})_{33}$	$(\mathbf{UB})_{33}$
38	5.1	11.8	26.3	0.14	0.36	0.7	1.8	4.3	9.4
39	3.6	8.2	18.8	0.10	0.22	0.50	1.1	2.7	5.9
40	2.8	6.5	14.5	0.073	0.16	0.37	0.84	1.8	4.3
50	0.95	2.3	5.2	0.016	0.040	0.085	0.11	0.29	0.62
60	0.66	1.7	3.6	0.0077	0.020	0.042	0.040	0.10	0.22
80	0.49	1.2	2.7	0.0032	0.0080	0.017	0.0093	0.023	0.051
100	0.44	1.1	2.4	0.0018	0.0045	0.010	0.0034	0.0085	0.019
150	0.35	0.88	2.0	0.00069	0.0017	0.0039	0.00058	0.0014	0.0033
200	0.33	0.84	1.8	0.00035	0.00091	0.0020	0.00016	0.00043	0.00094

Concerning $\mathbf{\Sigma}$ we randomly generated eigenvectors $\mathbf{\Gamma}$ via another covariance matrix and also randomly generated eigenvalues $\{\lambda_k\}$ uniformly on the interval $[2, 3.1]$. The eigenvalues build up a diagonal matrix $\mathbf{D} = (\lambda_k)$ and then the $\mathbf{\Sigma}$ which has been used in the simulations equals $\mathbf{\Sigma} = \mathbf{\Gamma D \Gamma}^\top$. Note that in the simulations $n = 40$, $r(\mathbf{C}) = 2$ and $q = 3$. The simulations were carried out for $p \in \{38, 39, 40, 50, 60, 80, 100, 150, 200\}$. According to Theorem 1 we have to assume that $p \geq 38$ and it can be shown that the theorem is not true for $p = 37$ and if $p < 37$ our bounds do not even exist. In (14) we have

$$E[(\mathbf{A}^\top\mathbf{S}^+\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{S}^+\mathbf{\Sigma}\mathbf{S}^+\mathbf{A}(\mathbf{A}^\top\mathbf{S}^+\mathbf{A})^{-1}] \tag{16}$$

which has to be estimated. The simulation data is generated according to the model $\mathbf{X}_i \sim N_{p,n}(\mathbf{ABC}, \mathbf{\Sigma}, \mathbf{I}_n)$, $i = 1, 2, \ldots, 500$, *i.e.*, there are 500 replicates performed in the simulation study. Let $\mathbf{S}_i^+$ denote the $\mathbf{S}^+$ from the ith simulation and we have

$$\mathbf{EST} = \frac{1}{n}\sum_{i=1}^{n} E[(\mathbf{A}^\top\mathbf{S}_i^+\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{S}_i^+\mathbf{\Sigma}\mathbf{S}_i^+\mathbf{A}(\mathbf{A}^\top\mathbf{S}_i^+\mathbf{A})^{-1}] \tag{17}$$

as an unbiased estimator of the expectation in (16). The results of the simulation study are presented in Table 1. In our case **EST** is of size 3×3. Moreover, we calculated the lower bound, **LB**, and the upper bound, **UB**, as

$$\mathbf{LB} = \lambda_p^3(\mathbf{\Sigma})\lambda_p(\mathbf{\Sigma}^{-1})\frac{p-1}{p-34}\frac{p-4}{35}(\mathbf{A}^\top\mathbf{\Sigma}\mathbf{A})^{-1}, \tag{18}$$

$$\mathbf{UB} = \lambda_1^3(\mathbf{\Sigma})\lambda_1(\mathbf{\Sigma}^{-1})\frac{p-1}{p-34}\frac{p-4}{35}(\mathbf{A}^\top\mathbf{\Sigma}\mathbf{A})^{-1} \tag{19}$$

which according to the theory should give upper and lower bounds of the expectation in (16). In Table 1 the diagonal elements of **EST**, **LB** and **UB** are presented. The results follow the theory, *i.e.*, $(\mathbf{LB})_{ii} < (\mathbf{EST})_{ii} < (\mathbf{UB})_{ii}$, $i \in \{1, 2, 3\}$. Moreover, when p increases $(\mathbf{EST})_{ii}$ becomes smaller and the difference $(\mathbf{UB})_{ii} - (\mathbf{LB})_{ii}$ is largest when $p = 38$. Thus, the results of Theorem 1 are in full agreement with the simulation study.

Acknowledgements

Dietrich von Rosen is supported by the Swedish Research Council (2017-03003).

References

Imori, S. and von Rosen, D. (2020). On the mean and dispersion of the Moore-Penrose generalized inverse of a Wishart matrix. Accepted by *Electronic Journal of Linear Algebra*.

Kollo, T. and von Rosen, D. (2005). *Advanced Multivariate Statistics with Matrices. Series: Mathematics and Its Applications*, **579**, Springer, Dordrecht.

Kollo, T., von Rosen, T. and von Rosen, D. (2011). Estimation in high-dimensional analysis and multivariate linear models. *Communications in Statistics: Theory Methods*, **40**, 1241–1253.

Potthoff, R. F. and Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika*, **51**, 313–326.

von Rosen, D. (2018). *Bilinear Regression Analysis. An Introduction. Lecture Notes in Statistics*, **220**. Springer, Cham.

Srivastava, M. S. and Khatri, C. G. (1979). *An Introduction to Multivariate Statistics.* North-Holland, New York-Oxford.