Multivariate Linear Models with Kronecker Product Structure

Muni Srivastava[1], Tatjana Nahtman[2] and Dietrich von Rosen[3]

[1] Department of Statistics, University of Toronto

[2] Institute of Mathematical Statistics, University of Tartu, Estonia

[3] Department of Biometry and Egineering, Swedish University of Agricultural Sciences

Outline

- Introduction

Outline

- Introduction
- Covariance structure of Kronecker type

Outline

- Introduction
- Covariance structure of Kronecker type
- Overparameterization

Outline

- Introduction
- Covariance structure of Kronecker type
- Overparameterization
- Growth Curve Model

Outline

- Introduction
- Covariance structure of Kronecker type
- Overparameterization
- Growth Curve Model
- Estimators

MANOVA

Classical multivariate statistical analysis is usually based on a vector with correlated components, for example, $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

MANOVA

Classical multivariate statistical analysis is usually based on a vector with correlated components, for example, $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

The correlation may be due to time dependence, spacial dependence or some other underlying latent process which is not observable.

Covariance with Kronecker structure

In many datasets we may have two processes which generate dependency.

Covariance with Kronecker structure

In many datasets we may have two processes which generate dependency.

For example,

- in environmental sciences: when studying catchment areas we have both spacial and temporal correlations,
- in neurosciences: when evaluating PET-image voxels, voxels are also both temporally and spacially correlated,
- in array technology: many antigens are represented on slides with observations over time, i.e. we have correlations between antigens as well over time.

Covariance with Kronecker structure

The main goal is to extend the classical model $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Covariance with Kronecker structure

The main goal is to extend the classical model $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to $\boldsymbol{y} \sim N_{pq}(\boldsymbol{\mu}, \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})$,

Covariance with Kronecker structure

The main goal is to extend the classical model $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to $\boldsymbol{y} \sim N_{pq}(\boldsymbol{\mu}, \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})$,

where $\boldsymbol{y} : pq \times 1,\ \boldsymbol{\mu} : pq \times 1,\ \boldsymbol{\Psi} : q \times q,\ \boldsymbol{\Sigma} : p \times p$, and $\otimes$ denotes the Kronecker product.

Covariance with Kronecker structure

The main goal is to extend the classical model $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to $\boldsymbol{y} \sim N_{pq}(\boldsymbol{\mu}, \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})$,

where $\boldsymbol{y} : pq \times 1,\ \boldsymbol{\mu} : pq \times 1,\ \boldsymbol{\Psi} : q \times q,\ \boldsymbol{\Sigma} : p \times p$, and $\otimes$ denotes the Kronecker product.

Both $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$ are unknown but it will be supposed that they are positive definite.

Due to the Kronecker product structure, we may convert $\boldsymbol{y} : pq \times 1$ into a matrix $\boldsymbol{Y} : p \times q$ which is matrix normally distributed, i.e. $\boldsymbol{Y} \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, where now $\boldsymbol{\mu}$ is a $p \times q$ matrix.

Wilks (1946), Votaw (1948), Srivastava (1965), Olkin (1973), Arnold (1973), Boik (1991), Naik & Rao (2001), Chaganty & Naik (2002), Lu & Zimmerman (2005), Roy & Khattree (2005).

Covariance with Kronecker structure

We assume that we have n observations, $\boldsymbol{Y}_i \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, whereas in the classical case one usually has one observation matrix $\boldsymbol{Y} \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{I})$.

Covariance with Kronecker structure

We assume that we have n observations, $\boldsymbol{Y}_i \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, whereas in the classical case one usually has one observation matrix $\boldsymbol{Y} \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{I})$.

The dispersion matrix of a matrix $\boldsymbol{Y}_i$ is defined by a vectorized form, i.e. $D[\boldsymbol{Y}_i] = D[vec(\boldsymbol{Y}_i)]$, where vec is the usual vec-operator. In our models

$$D[\boldsymbol{Y}_i] = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}.$$

.

Covariance with Kronecker structure

We assume that we have n observations, $\boldsymbol{Y}_i \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, whereas in the classical case one usually has one observation matrix $\boldsymbol{Y} \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{I})$.

The dispersion matrix of a matrix $\boldsymbol{Y}_i$ is defined by a vectorized form, i.e. $D[\boldsymbol{Y}_i] = D[vec(\boldsymbol{Y}_i)]$, where vec is the usual vec-operator. In our models

$$D[\boldsymbol{Y}_i] = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}.$$

.

We shall exploit how the independent "matrix-observations" can be used to estimate $\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Psi}$ with and without certain bilinear structures on $\boldsymbol{\mu}$.

Covariance with Kronecker structure

For the interpretation it is noted that $\boldsymbol{\Psi} : q \times q$ describes the covariance structure between the columns of $\boldsymbol{Y}_i$.

Covariance with Kronecker structure

For the interpretation it is noted that $\mathbf{\Psi} : q \times q$ describes the covariance structure between the columns of $\boldsymbol{Y}_i$.

The covariance between the columns will be the same for each row of $\boldsymbol{Y}_i$.

Covariance with Kronecker structure

For the interpretation it is noted that $\boldsymbol{\Psi} : q \times q$ describes the covariance structure between the columns of $\boldsymbol{Y}_i$.

The covariance between the columns will be the same for each row of $\boldsymbol{Y}_i$.

The other covariance matrix $\boldsymbol{\Sigma} : p \times p$ describes the covariance between the rows in $\boldsymbol{Y}_i$ which will be the same for each column.

Covariance with Kronecker structure

For the interpretation it is noted that $\boldsymbol{\Psi} : q \times q$ describes the covariance structure between the columns of $\boldsymbol{Y}_i$.

The covariance between the columns will be the same for each row of $\boldsymbol{Y}_i$.

The other covariance matrix $\boldsymbol{\Sigma} : p \times p$ describes the covariance between the rows in $\boldsymbol{Y}_i$ which will be the same for each column.

The product $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ takes into consideration both $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$.

Covariance with Kronecker structure

For the interpretation it is noted that $\mathbf{\Psi} : q \times q$ describes the covariance structure between the columns of $\boldsymbol{Y}_i$.

The covariance between the columns will be the same for each row of $\boldsymbol{Y}_i$.

The other covariance matrix $\mathbf{\Sigma} : p \times p$ describes the covariance between the rows in $\boldsymbol{Y}_i$ which will be the same for each column.

The product $\mathbf{\Psi} \otimes \mathbf{\Sigma}$ takes into consideration both $\mathbf{\Psi}$ and $\mathbf{\Sigma}$.

Indeed, $\mathbf{\Psi} \otimes \mathbf{\Sigma}$ tells us that the overall covariance consists of the products of the covariances in $\mathbf{\Psi}$ and $\mathbf{\Sigma}$, respectively.

Covariance with Kronecker structure

We have

$$Cov[y_{kl}, y_{rs}] = \sigma_{kr}\psi_{ls},$$

where $\boldsymbol{Y}_i = (y_{kl}), \boldsymbol{\Sigma} = (\sigma_{kr})$ and $\boldsymbol{\Psi} = (\psi_{ls})$.

Covariance with Kronecker structure

We have

$$Cov[y_{kl}, y_{rs}] = \sigma_{kr}\psi_{ls},$$

where $\boldsymbol{Y}_i = (y_{kl}), \boldsymbol{\Sigma} = (\sigma_{kr})$ and $\boldsymbol{\Psi} = (\psi_{ls})$.

$\boldsymbol{\Sigma}$ may consist of the time-dependent covariances and $\boldsymbol{\Psi}$ takes care of the spacial correlation, or for the array data $\boldsymbol{\Psi}$ models the dependency between antigens and $\boldsymbol{\Sigma}$ represents the correlation over time.

Covariance with Kronecker structure

We have

$$Cov[y_{kl}, y_{rs}] = \sigma_{kr}\psi_{ls},$$

where $\boldsymbol{Y}_i = (y_{kl}), \boldsymbol{\Sigma} = (\sigma_{kr})$ and $\boldsymbol{\Psi} = (\psi_{ls})$.

$\boldsymbol{\Sigma}$ may consist of the time-dependent covariances and $\boldsymbol{\Psi}$ takes care of the spacial correlation, or for the array data $\boldsymbol{\Psi}$ models the dependency between antigens and $\boldsymbol{\Sigma}$ represents the correlation over time.

Note that the correlation of y_{kl} and y_{rs} equals

$$corr[y_{kl}, y_{rs}] = \frac{\sigma_{kr}}{\sqrt{\sigma_{kk}\sigma_{ll}}}\frac{\psi_{ls}}{\sqrt{\psi_{ll}\psi_{ss}}}.$$

.

Growth Curve model

Let us assume that the mean of $\boldsymbol{Y}_i$ follows a multilinear model, i.e.

$$E[\boldsymbol{Y}_i] = \boldsymbol{ABC},$$

where $\boldsymbol{A} : p \times r$ and $\boldsymbol{C} : s \times q$ are known design matrices.

This type of mean structure was introduced by Potthoff $\&$ Roy (1964).

Under the assumption that $\boldsymbol{\Psi} = \boldsymbol{I}$ (or $\boldsymbol{\Psi}$ known), i.e. we have independent columns in $\boldsymbol{Y}_i$ this will give us the well known Growth Curve model. For details and references connected to the model it is referred to Srivastava $\&$ Khatri (1979), Srivastava $\&$ von Rosen (1998) or Kollo $\&$ von Rosen (2005).

Growth Curve model

Observe that if the matrix $\boldsymbol{Y}_i : p \times q$ is $N_{p,q}(\boldsymbol{ABC}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ distributed we may form a new matrix

$$\boldsymbol{Y} = (\boldsymbol{Y}_1 : \boldsymbol{Y}_2 : \ldots : \boldsymbol{Y}_n),$$

which is

$$N_{p,qn}(\boldsymbol{AB}(\mathbf{1}_n' \otimes \boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n \otimes \boldsymbol{\Psi}),$$

where $\mathbf{1}_n$ is a vector of $\mathbf{1}$s of size n.

Covariance with Kronecker structure

Aims

- The aim is to present estimating equations for estimating $\boldsymbol{B}, \boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$,

Covariance with Kronecker structure

Aims

- The aim is to present estimating equations for estimating $\boldsymbol{B}, \boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$,
- to show how to estimate the parameters when $\boldsymbol{FBG} = \boldsymbol{0}$ holds for known matrices $\boldsymbol{F}$ and $\boldsymbol{G}$.

Covariance with Kronecker structure

Aims

- The aim is to present estimating equations for estimating $\boldsymbol{B}, \boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$,
- to show how to estimate the parameters when $\boldsymbol{FBG} = \mathbf{0}$ holds for known matrices $\boldsymbol{F}$ and $\boldsymbol{G}$.
- based on the MLEs to consider the likelihood ratio test for testing $H_0 : \boldsymbol{FBG} = \mathbf{0}$ versus $H_1 : \boldsymbol{FBG} \neq \mathbf{0}$.

MLEs when $\mathbf{\Psi}$ is known

When $\mathbf{\Psi}$ is known we have a situation which is almost identical to the classical Growth Curve model setup. The main difference is that now we have n matrix observations instead of 1, i.e. $N_{p,q}(\boldsymbol{ABC}, \mathbf{\Sigma}, \mathbf{\Psi})$.

MLEs when $\boldsymbol{\Psi}$ is known

When $\boldsymbol{\Psi}$ is known we have a situation which is almost identical to the classical Growth Curve model setup. The main difference is that now we have n matrix observations instead of 1, i.e. $N_{p,q}(\boldsymbol{ABC}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$.

Since $\boldsymbol{\Psi}$ is positive definite, we transform the data, i.e. $\boldsymbol{Y}_i = \boldsymbol{Y}_i \boldsymbol{\Psi}^{-1/2}$, where $\boldsymbol{\Psi}^{1/2}$ is a positive definite square root of $\boldsymbol{\Psi}$.

MLEs when $\boldsymbol{\Psi}$ is known

When $\boldsymbol{\Psi}$ is known we have a situation which is almost identical to the classical Growth Curve model setup. The main difference is that now we have n matrix observations instead of 1, i.e. $N_{p,q}(\boldsymbol{ABC}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$.

Since $\boldsymbol{\Psi}$ is positive definite, we transform the data, i.e. $\boldsymbol{Y}_i = \boldsymbol{Y}_i\boldsymbol{\Psi}^{-1/2}$, where $\boldsymbol{\Psi}^{1/2}$ is a positive definite square root of $\boldsymbol{\Psi}$.

Let $\boldsymbol{Y} = (\boldsymbol{Y}_1 : \boldsymbol{Y}_2 : \ldots : \boldsymbol{Y}_n) : p \times qn$, then

$$\boldsymbol{Y} \sim N_{p,qn}(\boldsymbol{AB}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}), \boldsymbol{\Sigma}, \boldsymbol{I}).$$

MLEs when $\boldsymbol{\Psi}$ is known

From results in Srivastava & Khatri (1979) or Kollo & von Rosen (2005) it follows directly that

$$
\begin{aligned}
n\widehat{\boldsymbol{B}} = (\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{Y}(\mathbf{1}_n \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}) \\
+(\boldsymbol{A}')^{\circ}\boldsymbol{Z}_1 + \boldsymbol{A}'\boldsymbol{Z}_2\boldsymbol{C}^{\circ\prime}, \\
\boldsymbol{S} = \boldsymbol{Y}(\boldsymbol{I} - n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1/2})\boldsymbol{Y}',
\end{aligned}
$$

MLEs when $\boldsymbol{\Psi}$ is known

From results in Srivastava & Khatri (1979) or Kollo & von Rosen (2005) it follows directly that

$$n\widehat{\boldsymbol{B}} = (\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{Y}(\mathbf{1}_n \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-})$$
$$+(\boldsymbol{A}')^{\circ}\boldsymbol{Z}_1 + \boldsymbol{A}'\boldsymbol{Z}_2\boldsymbol{C}^{\circ\prime},$$
$$\boldsymbol{S} = \boldsymbol{Y}(\boldsymbol{I} - n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1/2})\boldsymbol{Y}',$$

where $\boldsymbol{A}'^{\circ}$ and $\boldsymbol{C}^{\circ}$ are any matrices which generate $\mathcal{C}(\boldsymbol{A}')^{\perp}$ and $\mathcal{C}(\boldsymbol{C})^{\perp}$, i.e. the orthogonal complements of $\mathcal{C}(\boldsymbol{A}')$ and $\mathcal{C}(\boldsymbol{C})$, respectively, and $\mathcal{C}(\cdot)$ denotes the column vector space.

MLEs when $\mathbf{\Psi}$ is known

From results in Srivastava $\&$ Khatri (1979) or Kollo $\&$ von Rosen (2005) it follows directly that

$$
\begin{aligned}
n\widehat{\boldsymbol{B}} &= (\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{Y}(\mathbf{1}_n \otimes \mathbf{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\mathbf{\Psi}^{-1}\boldsymbol{C}')^{-}) \\
&\qquad + (\boldsymbol{A}')^{\circ}\boldsymbol{Z}_1 + \boldsymbol{A}'\boldsymbol{Z}_2\boldsymbol{C}^{\circ\prime}, \\
\boldsymbol{S} &= \boldsymbol{Y}(\boldsymbol{I} - n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \mathbf{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\mathbf{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\mathbf{\Psi}^{-1/2})\boldsymbol{Y}',
\end{aligned}
$$

where $\boldsymbol{A}'^{\circ}$ and $\boldsymbol{C}^{\circ}$ are any matrices which generate $\mathcal{C}(\boldsymbol{A}')^{\perp}$ and $\mathcal{C}(\boldsymbol{C})^{\perp}$, i.e. the orthogonal complements of $\mathcal{C}(\boldsymbol{A}')$ and $\mathcal{C}(\boldsymbol{C})$, respectively, and $\mathcal{C}(\cdot)$ denotes the column vector space.

Moreover, $^{-}$ denotes an arbitrary g-inverse, and $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ are arbitrary matrices of proper size.

MLEs when $\boldsymbol{\Psi}$ is known

Furthermore,

$$nq\widehat{\boldsymbol{\Sigma}} = (\boldsymbol{Y} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))(\boldsymbol{Y} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))'$$

$$= \boldsymbol{S} + n^{-1}\boldsymbol{S}\boldsymbol{A}^{\circ}(\boldsymbol{A}^{\circ\prime}\boldsymbol{S}\boldsymbol{A}^{\circ})^{-}\boldsymbol{A}^{\circ\prime}\boldsymbol{Y}(\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{\Psi}^{-1}\boldsymbol{C})^{-}\boldsymbol{C}'\boldsymbol{\Psi}^{-1/2})\boldsymbol{Y}'$$
$$\times \boldsymbol{A}^{\circ}(\boldsymbol{A}^{\circ\prime}\boldsymbol{S}\boldsymbol{A}^{\circ})^{-}\boldsymbol{A}^{\circ\prime}\boldsymbol{S}.$$

MLEs when $\boldsymbol{\Psi}$ is known

Furthermore,

$$nq\widehat{\boldsymbol{\Sigma}} = (\boldsymbol{Y} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))(\boldsymbol{Y} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))'$$
$$= \boldsymbol{S} + n^{-1}\boldsymbol{S}\boldsymbol{A}^{\circ}(\boldsymbol{A}^{\circ\prime}\boldsymbol{S}\boldsymbol{A}^{\circ})^{-}\boldsymbol{A}^{\circ\prime}\boldsymbol{Y}(\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{\Psi}^{-1}\boldsymbol{C})^{-}\boldsymbol{C}'\boldsymbol{\Psi}^{-1/2})\boldsymbol{Y}'$$
$$\times \boldsymbol{A}^{\circ}(\boldsymbol{A}^{\circ\prime}\boldsymbol{S}\boldsymbol{A}^{\circ})^{-}\boldsymbol{A}^{\circ\prime}\boldsymbol{S}.$$

If rank($\boldsymbol{A}$)=r and rank($\boldsymbol{C}$)=s then $\widehat{\boldsymbol{B}}$ is uniquely estimated, i.e.

$$n\widehat{\boldsymbol{B}} = (\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{Y}(\mathbf{1}_n \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-1}).$$

MLEs when $\boldsymbol{\Psi}$ is known

Furthermore,

$$nq\widehat{\boldsymbol{\Sigma}} = (\boldsymbol{Y} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))(\boldsymbol{Y} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))'$$

$$= \boldsymbol{S} + n^{-1}\boldsymbol{S}\boldsymbol{A}^{\circ}(\boldsymbol{A}^{\circ'}\boldsymbol{S}\boldsymbol{A}^{\circ})^{-}\boldsymbol{A}^{\circ'}\boldsymbol{Y}(\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{\Psi}^{-1}\boldsymbol{C})^{-}\boldsymbol{C}'\boldsymbol{\Psi}^{-1/2})\boldsymbol{Y}'$$
$$\times \boldsymbol{A}^{\circ}(\boldsymbol{A}^{\circ'}\boldsymbol{S}\boldsymbol{A}^{\circ})^{-}\boldsymbol{A}^{\circ'}\boldsymbol{S}.$$

If rank($\boldsymbol{A}$)=r and rank($\boldsymbol{C}$)=s then $\widehat{\boldsymbol{B}}$ is uniquely estimated, i.e.

$$n\widehat{\boldsymbol{B}} = (\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{Y}(\mathbf{1}_n \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-1}).$$

Note that $\widehat{\boldsymbol{\Sigma}}$ is always uniquely estimated.

MLEs when $\boldsymbol{\Psi}$ is known

Turning to the restrictions $\boldsymbol{FBG} = \mathbf{0}$ it is observed that these restrictions are equivalent to

$$\boldsymbol{B} = (\boldsymbol{F}')^{\circ}\theta_1 + \boldsymbol{F}'\theta_2\boldsymbol{G}^{\circ\prime},$$

where θ_1 and θ_2 may be regarded as new parameters. From Theorem 4.1.15 in Kollo $\&$ von Rosen (2005) it follows that

$$\widehat{\boldsymbol{B}} = (\boldsymbol{F}')^{\circ}\widehat{\theta_1} + \boldsymbol{F}'\widehat{\theta}_2\boldsymbol{G}^{\circ\prime},$$

MLEs when $\boldsymbol{\Psi}$ is known

Turning to the restrictions $\boldsymbol{FBG} = \mathbf{0}$ it is observed that these restrictions are equivalent to

$$\boldsymbol{B} = (\boldsymbol{F}')^{\circ}\theta_1 + \boldsymbol{F}'\theta_2\boldsymbol{G}^{\circ\prime},$$

where θ_1 and θ_2 may be regarded as new parameters. From Theorem 4.1.15 in Kollo & von Rosen (2005) it follows that

$$\widehat{\boldsymbol{B}} = (\boldsymbol{F}')^{\circ}\widehat{\theta}_1 + \boldsymbol{F}'\widehat{\theta}_2\boldsymbol{G}^{\circ\prime},$$

where

$$\begin{aligned}\widehat{\theta}_2 = {} & (\boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}_1'\boldsymbol{S}_2^{-1}\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{-}\boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}_1'\boldsymbol{S}_2^{-1}\boldsymbol{T}_1\boldsymbol{Y}\boldsymbol{G}^{\circ}(\mathbf{1}_n\otimes\boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'\boldsymbol{G}^{\circ})(\boldsymbol{G}^{\circ\prime}\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}'\boldsymbol{G}^{\circ})^{-} \\ & + (\boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}_1')^{\circ}\boldsymbol{Z}_{11} + \boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}_1'\boldsymbol{Z}_{12}(\boldsymbol{G}^{\circ\prime}\boldsymbol{C})^{\circ\prime}\end{aligned}$$

MLEs when $\boldsymbol{\Psi}$ is known

with
$$\boldsymbol{T}_1 = \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{F'})^{\circ}((\boldsymbol{F'})^{\circ'}\boldsymbol{A}'\boldsymbol{S}_1^{-1}\boldsymbol{A}(\boldsymbol{F'})^{\circ})^{-}(\boldsymbol{F'})^{\circ'}\boldsymbol{A}'\boldsymbol{S}_1^{-1},$$

MLEs when $\boldsymbol{\Psi}$ is known

with

$$\boldsymbol{S}_1 \quad = \boldsymbol{Y}(\boldsymbol{I} - n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1/2})\boldsymbol{Y}'$$

is assumed to be positive definite,

MLEs when $\boldsymbol{\Psi}$ is known

with

$$\begin{aligned} \boldsymbol{S}_2 &= \boldsymbol{S}_1 + \boldsymbol{T}_1\boldsymbol{Y}(n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1/2}) \\ &\times (\boldsymbol{I} - n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'\boldsymbol{G}^{\circ}(\boldsymbol{G}^{\circ\prime}\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}'\boldsymbol{G}^{\circ})^{-}\boldsymbol{G}^{\circ\prime}\boldsymbol{C}\boldsymbol{\Psi}^{-1/2}) \\ &\times (n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1/2})\boldsymbol{Y}'\boldsymbol{T}_1', \end{aligned}$$

MLEs when $\mathbf{\Psi}$ is known

with

$$\widehat{\theta}_1 = (\boldsymbol{F}'^{\circ'}\boldsymbol{A}'\boldsymbol{S}_1^{-1}\boldsymbol{A}(\boldsymbol{F}')^{\circ})^{-}\boldsymbol{F}'^{\circ'}\boldsymbol{A}'\boldsymbol{S}_1^{-1}(\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{F}'\widehat{\theta}_2\boldsymbol{G}^{\circ'}\boldsymbol{C}\mathbf{\Psi}^{-1/2})\mathbf{\Psi}^{-1/2}\boldsymbol{C}'$$
$$\times(\boldsymbol{C}\mathbf{\Psi}^{-1}\boldsymbol{C}')^{-} + (\boldsymbol{F}'^{\circ'}\boldsymbol{A})^{\circ'}\boldsymbol{Z}_{21} + \boldsymbol{F}'^{\circ'}\boldsymbol{A}'\boldsymbol{Z}_{22}\boldsymbol{C}^{\circ'},$$

where $\boldsymbol{Z}_{ij}$ are arbitrary matrices,

MLEs when $\boldsymbol{\Psi}$ is known

Furthermore,

$$nq\widehat{\boldsymbol{\Sigma}} = (\boldsymbol{Y} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))(\boldsymbol{Y} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))'.$$

Explicit estimators when $\boldsymbol{\Psi}$ is unknown

If $\boldsymbol{Y}_i \sim N_{p,q}(\boldsymbol{ABC}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ it follows that

$$D[\boldsymbol{Y}_i] = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}.$$

Explicit estimators when $\mathbf{\Psi}$ is unknown

If $\boldsymbol{Y}_i \sim N_{p,q}(\boldsymbol{ABC}, \mathbf{\Sigma}, \mathbf{\Psi})$ it follows that

$$D[\boldsymbol{Y}_i] = \mathbf{\Psi} \otimes \mathbf{\Sigma}.$$

Since $c\mathbf{\Psi} \otimes c^{-1}\mathbf{\Sigma}$ ➔ the parameterization can not be uniquely interpreted, i.e. the model is overparameterized.

Explicit estimators when $\mathbf{\Psi}$ is unknown

If $\boldsymbol{Y}_i \sim N_{p,q}(\boldsymbol{ABC}, \mathbf{\Sigma}, \mathbf{\Psi})$ it follows that

$$D[\boldsymbol{Y}_i] = \mathbf{\Psi} \otimes \mathbf{\Sigma}.$$

Since $c\mathbf{\Psi} \otimes c^{-1}\mathbf{\Sigma}$ ➔ the parameterization can not be uniquely interpreted, i.e. the model is overparameterized.

If $\mathbf{\Psi}$ is unknown we will suppose that $\psi_{qq} = 1$.

Explicit estimators when $\boldsymbol{\Psi}$ is unknown

If $\boldsymbol{Y}_i \sim N_{p,q}(\boldsymbol{ABC}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ it follows that

$$D[\boldsymbol{Y}_i] = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}.$$

Since $c\boldsymbol{\Psi} \otimes c^{-1}\boldsymbol{\Sigma}$ ➔ the parameterization can not be uniquely interpreted, i.e. the model is overparameterized.

If $\boldsymbol{\Psi}$ is unknown we will suppose that $\psi_{qq} = 1$.

If all diagonal elements in $\boldsymbol{\Psi}$ are assumed to equal 1, it is easy to produce some heuristic estimators.

Explicit estimators when $\mathbf{\Psi}$ is unknown

If $\boldsymbol{Y}_i \sim N_{p,q}(\boldsymbol{ABC}, \mathbf{\Sigma}, \mathbf{\Psi})$ it follows that

$$D[\boldsymbol{Y}_i] = \mathbf{\Psi} \otimes \mathbf{\Sigma}.$$

Since $c\mathbf{\Psi} \otimes c^{-1}\mathbf{\Sigma}$ ➔ the parameterization can not be uniquely interpreted, i.e. the model is overparameterized.

If $\mathbf{\Psi}$ is unknown we will suppose that $\psi_{qq} = 1$.

If all diagonal elements in $\mathbf{\Psi}$ are assumed to equal 1, it is easy to produce some heuristic estimators.

The main idea is to produce estimators of $\boldsymbol{B}$ and $\mathbf{\Sigma}$ by neglecting the dependency among columns. Thereafter the off-diagonal elements in $\mathbf{\Psi}$ are estimated.

Explicit estimators when $\boldsymbol{\Psi}$ is unknown

Theorem 3.1. Let $\boldsymbol{Y} \sim N_{p,qn}(\boldsymbol{AB}(\mathbf{1}_n' \otimes \boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n \otimes \boldsymbol{\Psi})$, where diag($\boldsymbol{\Psi}$)=$\boldsymbol{I}$. Unbiased estimators of $\boldsymbol{ABC}$ and $\boldsymbol{\Sigma}$ are given by

$$\begin{aligned} n\boldsymbol{A}\widehat{\boldsymbol{B}}\boldsymbol{C} &= \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{Y}(\mathbf{1}_n \otimes \boldsymbol{C}'(\boldsymbol{C}\boldsymbol{C}')^{-}\boldsymbol{C}), \\ q(n-1)\widehat{\boldsymbol{\Sigma}} &= \boldsymbol{S} = \boldsymbol{Y}(\mathbf{1}_n^{\circ}(\mathbf{1}_n^{\circ\prime}\mathbf{1}_n^{\circ})^{-}\mathbf{1}_n^{\circ\prime} \otimes \boldsymbol{I})\boldsymbol{Y}'. \end{aligned}$$

Explicit estimators when $\boldsymbol{\Psi}$ is unknown

Theorem 3.2. Let $\boldsymbol{Y} \sim N_{p,qn}(\boldsymbol{A}\boldsymbol{B}(\mathbf{1}_n' \otimes \boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n \otimes \boldsymbol{\Psi})$, where diag($\boldsymbol{\Psi}$)=$\boldsymbol{I}$. A consistent estimator of the unknown elements in $\boldsymbol{\Psi}$ is given by

$$\begin{aligned} \widehat{\psi}_{kl} &= n^{-1} tr(\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{Y}(\boldsymbol{I} \otimes \boldsymbol{e}_k) - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{e}_k)) \\ &\qquad \times (\boldsymbol{Y}(\boldsymbol{I} \otimes \boldsymbol{e}_l) - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{e}_l))'), \ k \neq l, \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{A}\widehat{\boldsymbol{B}}\boldsymbol{e}_k &= n^{-1}\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{Y}(\mathbf{1}_n \otimes \boldsymbol{C}'(\boldsymbol{C}\boldsymbol{C}')^{-}\boldsymbol{e}_k), \\ \widehat{\boldsymbol{\Sigma}} &= (qn)^{-1}\boldsymbol{S} \end{aligned}$$

and $\boldsymbol{S}$ is given in Theorem 3.1.

Explicit estimators when $\mathbf{\Psi}$ is unknown

Theorem 3.3. Let $\boldsymbol{Y} \sim N_{p,qn}(\boldsymbol{AB}(\mathbf{1}'_n \otimes \boldsymbol{C}), \mathbf{\Sigma}, \boldsymbol{I}_n\mathbf{\Psi})$, where diag($\mathbf{\Psi}$)=$\boldsymbol{I}$ and $\boldsymbol{FBG} = 0$. Unbiased estimators of $\boldsymbol{ABC}$ and $\mathbf{\Sigma}$

$$\begin{aligned} \widehat{\boldsymbol{ABC}} &= \boldsymbol{A}(\boldsymbol{F}')^{\circ}\widehat{\theta}_1\boldsymbol{C} + \boldsymbol{AF}'\widehat{\theta}_2\boldsymbol{G}^{\circ'}\boldsymbol{C}, \\ q(n-1)\widehat{\mathbf{\Sigma}} &= \boldsymbol{S} = \boldsymbol{Y}(\mathbf{1}_n^{\circ}(\mathbf{1}_n^{\circ'}\mathbf{1}_n^{\circ})^{-}\mathbf{1}_n^{\circ'} \otimes \boldsymbol{I})\boldsymbol{Y}', \end{aligned}$$

Explicit estimators when $\boldsymbol{\Psi}$ is unknown

Theorem 3.3. Let $\boldsymbol{Y} \sim N_{p,qn}(\boldsymbol{AB}(\mathbf{1}_n' \otimes \boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n\boldsymbol{\Psi})$, where diag($\boldsymbol{\Psi}$)=$\boldsymbol{I}$ and $\boldsymbol{FBG} = 0$. Unbiased estimators of $\boldsymbol{ABC}$ and $\boldsymbol{\Sigma}$

$$\begin{aligned} \widehat{\boldsymbol{ABC}} &= \boldsymbol{A}(\boldsymbol{F}')^\circ \widehat{\theta}_1 \boldsymbol{C} + \boldsymbol{AF}'\widehat{\theta}_2 \boldsymbol{G}^{\circ'}\boldsymbol{C}, \\ q(n-1)\widehat{\boldsymbol{\Sigma}} &= \boldsymbol{S} = \boldsymbol{Y}(\mathbf{1}_n^\circ(\mathbf{1}_n^{\circ'}\mathbf{1}_n^\circ)^-\mathbf{1}_n^{\circ'} \otimes \boldsymbol{I})\boldsymbol{Y}', \end{aligned}$$

where

$$\begin{aligned} n\widehat{\theta}_2 = {} & (\boldsymbol{FA}'\boldsymbol{T}_1'\boldsymbol{S}^{-1}\boldsymbol{T}_1\boldsymbol{AF}')^-\boldsymbol{FA}'\boldsymbol{T}_1'\boldsymbol{S}^{-1}\boldsymbol{T}_1\boldsymbol{Y}(\mathbf{1}_n \otimes \boldsymbol{C}'\boldsymbol{G}^\circ(\boldsymbol{G}^{\circ'}\boldsymbol{CC}'\boldsymbol{G}^\circ)^-) \\ & + (\boldsymbol{FA}'\boldsymbol{T}_1')^\circ \boldsymbol{Z}_{11} + \boldsymbol{FA}'\boldsymbol{T}_1'\boldsymbol{Z}_{12}(\boldsymbol{G}^{\circ'}\boldsymbol{C})^{\circ'}, \\ \boldsymbol{T}_1 = {} & \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{F}')^\circ((\boldsymbol{F}')^{\circ'}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A}(\boldsymbol{F}')^\circ)^-(\boldsymbol{F}')^{\circ'}\boldsymbol{A}'\boldsymbol{S}^{-1}, \\ n\widehat{\theta}_1 = {} & ((\boldsymbol{F}')^{\circ'}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A}(\boldsymbol{F}')^\circ)^-\boldsymbol{F}'^{\circ'}\boldsymbol{A}'\boldsymbol{S}^{-1}(\boldsymbol{Y} - \boldsymbol{AF}'\widehat{\theta}_2\boldsymbol{G}^{\circ'}\boldsymbol{C}\boldsymbol{\Psi}^{-1/2})\boldsymbol{\Psi}^{-1/2}\boldsymbol{C}' \\ & \times \; (\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^- + (\boldsymbol{F}'^{\circ'}\boldsymbol{A})^{\circ'}\boldsymbol{Z}_{21} + \boldsymbol{F}'^{\circ'}\boldsymbol{A}'\boldsymbol{Z}_{22}\boldsymbol{C}^{\circ'}, \end{aligned}$$

where $\boldsymbol{Z}_{ij}$ are arbitrary matrices.

MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

The aim is to find maximum likelihood estimators of the parameters in the model

$$\boldsymbol{Y} \sim N_{p,qn}(\boldsymbol{A}\boldsymbol{B}(\mathbf{1}_n' \otimes \boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n \otimes \boldsymbol{\Psi}).$$

We assume that the uniqueness condition $\psi_{qq} = 1$ holds.

The other diagonal elements of $\boldsymbol{\Psi}$ will be positive but unknown (earlier we assumed $\mathrm{diag}(\boldsymbol{\Psi}) = bI$).

MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

The aim is to find maximum likelihood estimators of the parameters in the model

$$\boldsymbol{Y} \sim N_{p,qn}(\boldsymbol{AB}(\mathbf{1}_n' \otimes \boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n \otimes \boldsymbol{\Psi}).$$

We assume that the uniqueness condition $\psi_{qq} = 1$ holds.

The other diagonal elements of $\boldsymbol{\Psi}$ will be positive but unknown (earlier we assumed $\text{diag}(\boldsymbol{\Psi}) = bI$).

First we study the following model

$$\boldsymbol{Y} \sim N_{p,qn}(\boldsymbol{\mu}(\mathbf{1}_n' \otimes \boldsymbol{I}), \boldsymbol{\Sigma}, \boldsymbol{I} \otimes \boldsymbol{\Psi}).$$

and thereafter the model with $E[\boldsymbol{Y}] = \boldsymbol{AB}(\mathbf{1}_{\boldsymbol{n}}' \otimes \mathbf{C})$.

MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

The likelihood equals

$$L = c|\boldsymbol{\Sigma}|^{-1/2qn}|\boldsymbol{\Psi}|^{-1/2np}e^{-\frac{1}{2}tr\{\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\mathbf{1}_n'\otimes\boldsymbol{I}_q))(\boldsymbol{I}\otimes\boldsymbol{\Psi})^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\mathbf{1}_n'\otimes\boldsymbol{I}_q))'\}},$$

where c is a proportionality constant which does not depend on the parameters.

MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

The likelihood equals

$$L = c|\boldsymbol{\Sigma}|^{-1/2qn}|\boldsymbol{\Psi}|^{-1/2np}e^{-\frac{1}{2}tr\{\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\mathbf{1}_n'\otimes\boldsymbol{I}_q))(\boldsymbol{I}\otimes\boldsymbol{\Psi})^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\mathbf{1}_n'\otimes\boldsymbol{I}_q))'\}},$$

where c is a proportionality constant which does not depend on the parameters.

The MLE of the mean equals

$$\widehat{\boldsymbol{\mu}} = n^{-1}\boldsymbol{Y}(\mathbf{1}_n \otimes \boldsymbol{I}_q)$$

MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

The MLE of $\boldsymbol{\Sigma}$ equals

$$nq\widehat{\boldsymbol{\Sigma}} = \boldsymbol{Y}(\mathbf{1}_n^o(\mathbf{1}_n^{o\prime}\mathbf{1}_n^o)\mathbf{1}_n^{o\prime} \otimes \boldsymbol{\Psi})\boldsymbol{Y}'.$$

MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

The MLE of $\boldsymbol{\Sigma}$ equals

$$nq\widehat{\boldsymbol{\Sigma}} = \boldsymbol{Y}(\mathbf{1}_n^o(\mathbf{1}_n^{o'}\mathbf{1}_n^o)\mathbf{1}_n^{o'} \otimes \boldsymbol{\Psi})\boldsymbol{Y}'.$$

When estimating $\boldsymbol{\Psi}$ we have to take into account that $\psi_{qq} = 1$.

MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

The MLE of $\boldsymbol{\Sigma}$ equals

$$nq\widehat{\boldsymbol{\Sigma}} = \boldsymbol{Y}(\mathbf{1}_n^o(\mathbf{1}_n^{o\prime}\mathbf{1}_n^o)\mathbf{1}_n^{o\prime} \otimes \boldsymbol{\Psi})\boldsymbol{Y}'.$$

When estimating $\boldsymbol{\Psi}$ we have to take into account that $\psi_{qq} = 1$. The idea is to condition with respect to columns $\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{iq}$. Put

$$\mathbf{Z}_{ik} = (\boldsymbol{y}_{ik} : \boldsymbol{y}_{ik+1} : \cdots : \boldsymbol{y}_{iq})$$

and let $f_{\boldsymbol{Y}_i}$ denote the density function for $\boldsymbol{Y}_i$. Then,

MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

The MLE of $\boldsymbol{\Sigma}$ equals

$$nq\widehat{\boldsymbol{\Sigma}} = \boldsymbol{Y}(\mathbf{1}_n^o(\mathbf{1}_n^{o\prime}\mathbf{1}_n^o)\mathbf{1}_n^{o\prime} \otimes \boldsymbol{\Psi})\boldsymbol{Y}'.$$

When estimating $\boldsymbol{\Psi}$ we have to take into account that $\psi_{qq} = 1$. The idea is to condition with respect to columns $\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{iq}$. Put

$$\mathbf{Z}_{ik} = (\boldsymbol{y}_{ik} : \boldsymbol{y}_{ik+1} : \cdots : \boldsymbol{y}_{iq})$$

and let $f_{\boldsymbol{Y}_i}$ denote the density function for $\boldsymbol{Y}_i$. Then,

$$f_{\boldsymbol{Y}_i} = \prod_{k=2}^{q} (f_{\boldsymbol{y}_{ik-1}|\mathbf{Z}_{ik}}) f_{\boldsymbol{y}_{iq}}.$$

MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

The conditional distributions are normal with parameters

$$\mathbf{B}^k = (\boldsymbol{\Psi}_{22}^k)^{-1}\boldsymbol{\Psi}_{21}^k$$
$$\psi_{1\bullet 2}^k = \psi_{11}^k - \boldsymbol{\Psi}_{12}^k(\boldsymbol{\Psi}_{22}^k)^{-1}\boldsymbol{\Psi}_{21}^k,$$

where

$$\boldsymbol{\Psi}^k = (0 : \mathbf{I}_t)\boldsymbol{\Psi}\begin{pmatrix} 0 \\ \mathbf{I}_t \end{pmatrix}, \quad t = q - k - 2,$$

$$\boldsymbol{\Psi}^k = \begin{pmatrix} \psi_{11}^k & \boldsymbol{\Psi}_{12}^k \\ \boldsymbol{\Psi}_{21}^k & \boldsymbol{\Psi}_{22}^k \end{pmatrix}, \qquad \begin{pmatrix} 1 \times 1 & 1 \times t-1 \\ t-1 \times 1 & t-1 \times t-1 \end{pmatrix}$$

MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

The parameters $\{\mathbf{B}^k, \boldsymbol{\Psi}^k_{1\bullet 2}\}$ are in one-to-one correspondence with $\boldsymbol{\Psi}$ if $\psi_{qq} = 1$.

MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

The parameters $\{\mathbf{B}^k, \boldsymbol{\Psi}^k_{1\bullet 2}\}$ are in one-to-one correspondence with $\boldsymbol{\Psi}$ if $\psi_{qq} = 1$. Estimation equations are given by

$$\begin{aligned}
&\widehat{\boldsymbol{B}}^k \\
&= (\sum_{i=1}^{n}(\boldsymbol{z}_{ik} - \boldsymbol{\mu}_2^k)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{z}_{ik} - \boldsymbol{\mu}_2^k))^{-1}\sum_{i=1}^{n}(\boldsymbol{z}_{ik} - \boldsymbol{\mu}_2^k)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_{k-1} - \boldsymbol{\mu}_1^k), \\
&k = 2, 3, \ldots, q, \\
&np\widehat{\boldsymbol{\Psi}}^k_{1\bullet 2} = tr\{\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_{k-1} - (\boldsymbol{\mu}_1^k + (\boldsymbol{z}_k - (\mathbf{1}_n' \otimes \boldsymbol{\mu}_2^k))(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{B}}^k))) \\
&\times(\boldsymbol{y}_{k-1} - (\boldsymbol{\mu}_1^k + (\boldsymbol{z}_k - (\mathbf{1}_n' \otimes \boldsymbol{\mu}_2^k))(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{B}}^k)))'\}.
\end{aligned}$$

where

$$\boldsymbol{\mu}^k = \boldsymbol{\mu}\begin{pmatrix} 0 \\ \mathbf{I}_t \end{pmatrix}$$

MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

Estimation equations when $\psi_{qq} = 1$; Summary.

$$\boldsymbol{\mu} = n^{-1}\boldsymbol{Y}(\mathbf{1}_n \otimes \boldsymbol{I}_q),$$

$$nq\boldsymbol{\Sigma} = \boldsymbol{Y}(\mathbf{1}_n^{\circ}(\mathbf{1}_n^{\circ\prime}\mathbf{1}_n^{\circ})^{-}\mathbf{1}_n^{\circ\prime} \otimes \boldsymbol{\Psi}^{-1})\boldsymbol{Y}',$$

$$\boldsymbol{B}^k = (\sum_{i=1}^{n}(\boldsymbol{z}_{ik} - \boldsymbol{\mu}_2^k)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{z}_{ik} - \boldsymbol{\mu}_2^k))^{-1}\sum_{i=1}^{n}(\boldsymbol{z}_{ik} - \boldsymbol{\mu}_2^k)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_{k-1} - \boldsymbol{\mu}_1^k),$$

$$np\boldsymbol{\Psi}_{1\bullet 2}^k = tr\{\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_{k-1} - (\boldsymbol{\mu}_1^k + (\boldsymbol{z}_k - (\mathbf{1}_n' \otimes \boldsymbol{\mu}_2^k))(\boldsymbol{I}_n \otimes \boldsymbol{B}^k)))()'\},$$

MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

Estimation equations when $\psi_{qq} = 1$ can also be obtained when $\mu = \mathbf{ABC}$. In this case we replace the equation concerning μ by an equation for $\mathbf{B}$ which is obtained from a ML-approach where $\boldsymbol{\Psi}$ is known:

$$n\widehat{\boldsymbol{B}} = (\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{Y}(\mathbf{1}_n \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-1}).$$

$$nq\widehat{\boldsymbol{\Sigma}} = (\boldsymbol{Y} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))(\boldsymbol{Y} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))'$$

Thank you!