Explicit Estimators for a Banded Covariance Matrix in a Multivariate Normal Distribution

Presentation

Authors: Emil Karlsson, Martin Singull

Matematiska institutionen
Linköpings universitet

August 25, 2014

History

- Patterned covariance matrices
- Banded covariance matrices
- Methods: explicit, maximum likelihood and back again

List of symbols

$\boldsymbol{A}_{m,n}$ - matrix of size $m \times n$
$M_{m,n}$ - the set of all matrices of size $m \times n$
a_{ij} - matrix element of the i-th row and j-th column
$\boldsymbol{a}_n$ - vector of size n
c - scalar
$\boldsymbol{X}$ - random matrix
$\boldsymbol{x}$ - random vector
X - random variable

Proposition 1

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

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

$$\hat{\sigma}_{i,i} = \frac{1}{n}\boldsymbol{x}_i'\boldsymbol{C}\boldsymbol{x}_i \textit{ for } i = 1,\ldots,p,$$

$$\hat{\sigma}_{i,i+1} = \frac{1}{n}\hat{\boldsymbol{r}}_i'\boldsymbol{C}\boldsymbol{x}_{i+1} \textit{ for } i = 1,\ldots,p-1,$$

$$\textit{where } \hat{\boldsymbol{r}}_1 = \boldsymbol{y}_1 \textit{ and } \hat{\boldsymbol{r}}_i = \boldsymbol{x}_i - \hat{s}_i\hat{\boldsymbol{r}}_{i-1} \textit{ for } i = 2,\ldots,p-1,$$

$$\hat{s}_i = \frac{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{C}\boldsymbol{x}_i}{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{C}\boldsymbol{x}_{i-1}},$$

$$\textit{where } \boldsymbol{C} = \boldsymbol{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'.$$

Theorem 1

*The estimator $\hat{\boldsymbol{\mu}} = (\hat{\mu}_1, \ldots, \hat{\mu}_p)'$ given in Proposition 1 is unbiased and consistent, and the estimator $\hat{\boldsymbol{\Sigma}}_{(p)}^{(m)} = (\hat{\sigma}_{ij})$ is consistent.[**?**]*

Goals:

- Find an unbiased estimator for the covariance matrix.
- Generalize results into a general linear model.

Limitations:

- Study the case where $\mathbf{\Sigma}_{(1)}^{(p)}$ instead of $\mathbf{\Sigma}_{(m)}^{(p)}$.

Proposition 2

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

$$\ldots$$

$$\hat{\sigma}_{i,i+1} = \frac{1}{n}\boldsymbol{x}'_i\boldsymbol{A}_{i-1}\boldsymbol{x}_{i+1} \textit{ for } i = 1, \ldots, p-1,$$

$$\textit{where } \boldsymbol{A}_i = \boldsymbol{C} - \boldsymbol{C}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}'_i\boldsymbol{C}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}'_i\boldsymbol{C},$$

$$\textit{with } \boldsymbol{A}_0 = \boldsymbol{C},$$

$$\textit{where } \hat{\boldsymbol{r}}_1 = \boldsymbol{y}_1 \textit{ and } \hat{\boldsymbol{r}}_i = \boldsymbol{x}_i - \frac{\hat{\boldsymbol{r}}'_{i-1}\boldsymbol{C}\boldsymbol{x}_i}{\hat{\boldsymbol{r}}'_{i-1}\boldsymbol{C}\boldsymbol{x}_{i-1}}\hat{\boldsymbol{r}}_{i-1} \textit{ for } i = 2, \ldots, p-1,$$

$$\textit{with } \boldsymbol{C} = \boldsymbol{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}'_n.$$

Definition 1

Let $\boldsymbol{x} \sim N_n(\boldsymbol{\mu_x}, \boldsymbol{I}_n)$,$\boldsymbol{y} \sim N_n(\boldsymbol{\mu_y}, \boldsymbol{I}_n)$ *and* $\boldsymbol{A} \in M_{n,n}$*. Then* $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{y}$ *is called a bilinear form.*

Theorem 2

The bilinear form $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{y}$ *has the following properties.*

(i) $E[\boldsymbol{x}'\boldsymbol{A}\boldsymbol{y}] = tr(\boldsymbol{A}\,\mathrm{cov}(\boldsymbol{x}, \boldsymbol{y}))$

(ii) $\mathrm{var}[\boldsymbol{x}'\boldsymbol{A}\boldsymbol{y}] = tr(\boldsymbol{A}\,\mathrm{cov}(\boldsymbol{x}, \boldsymbol{y}))^2 + tr(\boldsymbol{A}\,\mathrm{var}(\boldsymbol{x})\boldsymbol{A}\,\mathrm{var}(\boldsymbol{y})) = tr(\boldsymbol{A})\,\mathrm{cov}(\boldsymbol{x}, \boldsymbol{y})^2 + tr(\boldsymbol{A}^2)\,\mathrm{var}(\boldsymbol{x})\,\mathrm{var}(\boldsymbol{y})$.

The central matrix for $\hat{\sigma}_{i,i+1} = \frac{1}{n}\boldsymbol{x}_i'\boldsymbol{A}_{i-1}\boldsymbol{x}_{i+1}$

$$\boldsymbol{A}_i = \boldsymbol{C} - \boldsymbol{C}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{C}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{C}$$

Properties:

- Idempotent, $\boldsymbol{A}_i^2 = \boldsymbol{A}_i$
- Symmetric, $\boldsymbol{A}_i' = \boldsymbol{A}_i$

Proposition 3

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

$$\hat{\sigma}_{ii} = \frac{1}{n-1}\boldsymbol{x}_i'\boldsymbol{C}\boldsymbol{x}_i \textit{ for } i = 1, \dots, p,$$

$$\hat{\sigma}_{12} = \frac{1}{n-1}\boldsymbol{x}_1'\boldsymbol{C}\boldsymbol{x}_2,$$

$$\hat{\sigma}_{i,i+1} = \frac{1}{n-2}\boldsymbol{x}_i'\boldsymbol{A}_{i-1}\boldsymbol{x}_{i+1} \textit{ for } i = 2, \dots, p-1,$$

$$\textit{where } \boldsymbol{A}_i = \boldsymbol{C} - \boldsymbol{C}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{C}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{C}, \textit{ with } \boldsymbol{A}_0 = \boldsymbol{C},$$

$$\textit{where } \hat{\boldsymbol{r}}_1 = \boldsymbol{y}_1 \textit{ and } \hat{\boldsymbol{r}}_i = \boldsymbol{x}_i - \frac{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{C}\boldsymbol{x}_i}{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{C}\boldsymbol{x}_{i-1}}\hat{\boldsymbol{r}}_{i-1} \textit{ for } i = 2, \dots, p-1,$$

$$\textit{with } \boldsymbol{C} = \boldsymbol{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'.$$

Theorem 3

The estimators from Proposition 3 are unbiased and consistent.

Variance is known:

- $var(\hat{\sigma}_{i,i+1}) = \frac{\sigma_{i,i+1}^2 + \sigma_{ii}\sigma_{i+1,i+1}}{n-2}$

Assumptions:

General linear model

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{E} \sim N_{n,p}(\boldsymbol{X}\boldsymbol{B}, \boldsymbol{I}_n, \boldsymbol{\Sigma}_{(1)}^{(p)})$$

- $\boldsymbol{Y}$ and $\boldsymbol{E}$ are $n \times m$ random matrices
- $\boldsymbol{X}$ is a known $n \times p$-design matrix with full rank
- $\boldsymbol{B}$ is an unknown $p \times m$-matrix of regression coefficients.
- $n \geq m + p$, and the rows of the error matrix $\boldsymbol{E}$ are independent $N_m(\mathbf{0}, \boldsymbol{\Sigma})$ random vectors.

Problems when generalizing

- The expected value $\boldsymbol{XB}$ differs from $\mu\mathbf{1}'$.
- The design matrix affects the degrees of freedom.

Transformation:

$(\boldsymbol{Y} - \boldsymbol{XB}) \sim N(\mathbf{0}, \boldsymbol{I}_n, \boldsymbol{\Sigma})$ can be treated as

$(\boldsymbol{y}_i - \boldsymbol{X}\boldsymbol{b}_i) \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ were each part is handled separately

Proposition 4

Let $\boldsymbol{Y} = \boldsymbol{XB} \sim N_{p,n}(\boldsymbol{XB}, \boldsymbol{\Sigma}_{(p)}^{(1)}, \boldsymbol{I}_n)$, where rank$(\boldsymbol{X}) = k$. Explicit estimators are given by

$$\hat{\boldsymbol{B}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y},$$

$$\hat{\sigma}_{ii} = \frac{1}{n-k}\boldsymbol{y}_i'\boldsymbol{D}\boldsymbol{y}_i \text{ for } i = 1, \ldots, p,$$

$$\hat{\sigma}_{i,i+1} = \frac{1}{n-k-1}\boldsymbol{y}_i'\boldsymbol{A}_{i-1}\boldsymbol{x}_{i+1} \text{ for } i = 2, \ldots, p-1,$$

$$\text{where } \boldsymbol{A}_i = \boldsymbol{D} - \boldsymbol{D}\hat{\boldsymbol{r}}_i(\hat{\boldsymbol{r}}_i'\boldsymbol{D}\hat{\boldsymbol{r}}_i)^{-1}\hat{\boldsymbol{r}}_i'\boldsymbol{D} \text{ with } \boldsymbol{A}_0 = \boldsymbol{D},$$

$$\text{where } \hat{\boldsymbol{r}}_1 = \boldsymbol{y}_1 \text{ and } \hat{\boldsymbol{r}}_i = \boldsymbol{y}_i - \frac{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{D}\boldsymbol{y}_i}{\hat{\boldsymbol{r}}_{i-1}'\boldsymbol{D}\boldsymbol{y}_{i-1}}\hat{\boldsymbol{r}}_{i-1} \text{ for } i = 2, \ldots, p-1,$$

$$\text{with } \boldsymbol{D} = \boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'.$$

Theorem 4

The estimators from Proposition 4 are unbiased and consistent.

Known variance:

$$var(\hat{\sigma}_{i,i+1}) = \frac{1}{n-2}(\boldsymbol{A})(\sigma_{i,i+1}^2 + \sigma_{i,i}\sigma_{i+1,i+1})$$

Based on the 100000 averages of samples with n=20, explicit unbiased average estimators, with true value within parenthesis, are given by,

$$\hat{\boldsymbol{\Sigma}}_{new} = \begin{pmatrix} 4.99501(5) & 1.99590(2) & 0.00000 & 0.00000 \\ 1.99590(2) & 4.99238(5) & 0.99678(1) & 0.00000 \\ 0.00000 & 0.99678(1) & 5.00026(5) & 3.00265(3) \\ 0.00000 & 0.00000 & 3.00265(3) & 5.00368(5) \end{pmatrix},$$

and the previous estimators are given by,

$$\hat{\boldsymbol{\Sigma}}_{prev} = \begin{pmatrix} 4.74526(5) & 1.89611(2) & 0.00000 & 0.00000 \\ 1.89611(2) & 4.74276(5) & 0.89710(1) & 0.00000 \\ 0.00000 & 0.89710(1) & 4.75025(5) & 2.70239(3) \\ 0.00000 & 0.00000 & 2.70239(3) & 4.75350(5) \end{pmatrix}.$$

Based on the 100000 averages of samples with n=80 and 20 regression parameters, where $\boldsymbol{X}$ and $\boldsymbol{B}$ were randomly generated, unbiased explicit average estimators, with true value within parenthesis, are given by,

$$\hat{\boldsymbol{\Sigma}}_{new} = \begin{pmatrix} 3.9986(4) & 0.9997(1) & 0 & 0 & 0 \\ 0.9997(1) & 3.0051(3) & 2.0024(2) & 0 & 0 \\ 0 & 2.0024(2) & 4.9989(5) & 2.9976(3) & 0 \\ 0 & 0 & 2.9976(3) & 4.9941(5) & 2.9950(3) \\ 0 & 0 & 0 & 2.9950(3) & 4.9935(5) \end{pmatrix},$$

and the previous estimators are given by,

$$\hat{\boldsymbol{\Sigma}}_{prev} = \begin{pmatrix} 2.9989(4) & 0.7497(2) & 0 & 0 & 0 \\ 0.7497(2) & 2.2538(3) & 1.4768(2) & 0 & 0 \\ 0 & 1.4768(2) & 3.7492(5) & 2.2107(3) & 0 \\ 0 & 0 & 2.2107(3) & 3.7455(5) & 2.2088(3) \\ 0 & 0 & 0 & 2.2088(3) & 3.7451(5) \end{pmatrix}.$$

Conclusion:

- The unbiased version makes an considerable improvement

Topics:

- Find unbiased estimator for $\boldsymbol{\Sigma}_{(m)}^{(p)}$
- Compare it to other estimators(for example MLE) for banded matrices.
- Study the variance to determine efficiency