Spatio-temporal mixed linear models in Small Area Estimation

Tatjana von Rosen
Department of Statistics
Stockholm University
Sweden

Outline

Background

Abstract

- This work concerns small area estimation from longitudinal surveys where data exhibit spatio-temporal patterns.

Abstract

- This work concerns small area estimation from longitudinal surveys where data exhibit spatio-temporal patterns.
- Area-level mixed linear model is proposed to take into account possible correlation among the neighboring areas and time points.

Abstract

- This work concerns small area estimation from longitudinal surveys where data exhibit spatio-temporal patterns.
- Area-level mixed linear model is proposed to take into account possible correlation among the neighboring areas and time points.
- The covariance structures suitable for describing spatio-temporal dependence are discussed.

Abstract

- This work concerns small area estimation from longitudinal surveys where data exhibit spatio-temporal patterns.
- Area-level mixed linear model is proposed to take into account possible correlation among the neighboring areas and time points.
- The covariance structures suitable for describing spatio-temporal dependence are discussed.

Introduction

➢Sample surveys provide a cost effective way of obtaining estimates for characteristics of interest at both population and subpopulation levels (small areas) which are not available in administrative registers.

➢In case of register-based statistics which comprise administrative data from registers and administrative systems, there is no problem to make regional breakdowns of data.

➢In theory, register-based statistics can be broken down to any level. The only limitation is that the statistics should not disclose individuals.

Introduction

➢Regarding statistics based on data from sample surveys, the problem is rather the opposite.

➢The risk of disclosure of individuals is practically non-existent but the ability to break down the statistics on small areas is much more difficult when the samples get smaller as the larger number of breakdowns is made.

Introduction

➢Regional statistics play an important role in the governmental decision making when distributing funds based on regional statistics concerning e.g. public health, criminality, unemployment, etc. Hence, reliable estimates are of utmost importance.

➢Small area estimation has received a lot of attention due to its applications in official statistics.

Introduction

➢Longitudinal data are usually collected in order to get information about changes over time. Due to a long tradition of official statistics and register data in the Nordic countries, longitudinal survey data is often available.

Introduction

➢Longitudinal data are usually collected in order to get information about changes over time. Due to a long tradition of official statistics and register data in the Nordic countries, longitudinal survey data is often available.

➢For example, victimization surveys have been conducted in Estonia in 1993, 1995, 2000, 2004 and 2009. Many small areas had a low number of respondents.

SAE

Small Area Estimation

➢Small area estimation is widely used for producing estimates of population parameters for areas (domains) with small, or even zero, sample sizes.

Small Area Estimation

➢Small area estimation is widely used for producing estimates of population parameters for areas (domains) with small, or even zero, sample sizes.

➢In the case of small domain sample sizes, estimation that only relies on domain-specific observations may lead to estimates with large variance.

Small Area Estimation

- ➢Small area estimation is widely used for producing estimates of population parameters for areas (domains) with small, or even zero, sample sizes.
- ➢In the case of small domain sample sizes, estimation that only relies on domain-specific observations may lead to estimates with large variance.
- ➢One possible solution is to employ estimation that borrows information from related small areas through statistical models using administrative data (registers), in order to increase precision of the estimates. Such estimation is often based on mixed linear models providing a link to a related small area through the use of supplementary data.

➢In SAE it is often assumed that (population) units in different small areas are uncorrelated.

➢However, in practice the boundaries that define a small area are arbitrarily set and there appears to be no good reason why population units that belong to neighbouring small areas should not be correlated.

Small Area Estimation

➢In SAE it is often assumed that (population) units in different small areas are uncorrelated.

➢However, in practice the boundaries that define a small area are arbitrarily set and there appears to be no good reason why population units that belong to neighbouring small areas should not be correlated.

➢For example, with agricultural, environmental, economic and epidemiological data, units that are spatially close may be more related than units that are further apart, although they may belong to different small areas.

➢It is therefore often reasonable to assume the correlation for the neighbouring areas.

Small Area Estimation

➢Mixed models have been frequently used in a various small area applications, since they offer great flexibility in combining information from various sources, in handling intra- and interarea correlations.

➢When longitudinal and cross-sectional data are available, MLM might be of use to take simultaneously advantage of spatial similarities among small areas and the temporal relationships of the data in order to improve the efficiency of the small area estimators.

Preliminaries

Mixed Linear Model

Linear mixed models are extensively used in many research areas due to the flexibility they offer for modelling longitudinal and spatial data.

All linear mixed models considered in this work can be viewed as special cases of the following mixed linear model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{\varepsilon},$$

$\boldsymbol{y}$ is an n-vector of observable random variables,
$\boldsymbol{\beta}$ is a p-vector of fixed effects,
$\boldsymbol{X} : n \times p$ and $\boldsymbol{Z} : n \times k$ are known design matrices,
$\boldsymbol{u} : k \times 1$ is a vector of random effects,
$\boldsymbol{\varepsilon} : n \times 1$ is a vector of random errors.
We suppose that $E(\boldsymbol{u}) = 0$, $E(\boldsymbol{\varepsilon}) = 0$ and

$$Var \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{\varepsilon} \end{pmatrix} = \begin{pmatrix} \boldsymbol{G} & 0 \\ 0 & \boldsymbol{R} \end{pmatrix}.$$

Hence, $\boldsymbol{V} = Var(\boldsymbol{y}) = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}^T + \boldsymbol{R}$.
Assuming normality, $\boldsymbol{y} \sim N(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{V})$.

The Best Linear Unbiased Estimator (BLUE) of fixed effects is given by

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{Y}.$$

Best Linear Unbiased Predictor (BLUP) of random effects is given by

$$\hat{\boldsymbol{u}} = \boldsymbol{G}\boldsymbol{Z}^T\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}).$$

Mixed Linear Models and SAE

- Small area refers to a small geographical area or a group for which little information is obtained from the sample survey.

When only a few observations are available from a given small area, the direct estimator based only on the data from the small area is likely to be unreliable.

- The key question of small area estimation is how to obtain reliable regional statistics when the sample data contain too few observations to assure adequate precision for statistical inference.

- Nowadays, a common solution is to use statistical models which make it possible to borrow strength for the estimation by utilizing data from similar or neighboring areas, or from similar surveys conducted earlier, i.e. borrowing strength over space or/and time.
- Moreover, these models make use of the auxiliary variables that might be available from administrative records or censuses.

Mixed Linear Models and SAE

Let θ_i be the parameter of interest (some function of the small area mean and $\hat{\theta}_i$ be the direct estimator of θ_i (survey-based estimate), $i = 1, \ldots, m$.

Assume that auxiliary data are available at area level, i.e. we have area-specific data vectors $\boldsymbol{x}_i = (x_{1i}, \ldots, x_{pi})$ with known values for each area. A design model (sampling model) can be expressed as following:

$$\hat{\theta}_i = \theta_i + \varepsilon_i,$$

where the ε_i's are independent sampling errors with zero mean and known sampling variances σ_i^2.

Mixed Linear Models and SAE

The key assumption is that θ_i is related to the area-specific auxiliary data through a linear model (linking model):

$$\theta_i = \boldsymbol{x}_i'\boldsymbol{\beta} + u_i,$$

where $u_i \sim N(0, \sigma_u^2)$, $i = 1, \dots, m$. Combining the linking model with the sampling model yields the following mixed linear model:

$$\hat{\theta}_i = \boldsymbol{x}_i'\boldsymbol{\beta} + u_i + \varepsilon_i,$$

where $\boldsymbol{\beta} : p \times 1$ defines the effects of the auxiliary variables, $\boldsymbol{x}_i : p \times 1$ is a vector of known constants, u_i is area-specific random effects, and ε_i is the sampling error.

Examples

The nested error regression model (individual level model):

$$y_{ij} = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + u_i + \varepsilon_{ij},$$

where $i = 1, \ldots, m$, $j = 1, \ldots, n_i$, k is the number of small areas, $N = \sum_{i=1}^{m} n_i$, $\boldsymbol{x}_{ij} : p \times 1$ is the vector of explanatory variables, $\boldsymbol{\beta} : p \times 1$ is an unknown vector of regression coefficients, and u_i's and ε_{ij}'s are mutually independently distributed, $u_i \sim N(0, \sigma_u^2)$ and $\varepsilon_{ij} \sim N(0, \sigma^2)$, respectively.

In matrix notation, this model can be expressed as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{\varepsilon}.$$

•Fay-Herriot model (area level model):

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + u_i + \varepsilon_i,$$

where $i = 1, \ldots, m$, m is the number of small areas, $\boldsymbol{x}_i : p \times 1$ is the vector of explanatory variables, $\boldsymbol{\beta} : p \times 1$ is an unknown vector of regression coefficients, u_i's and ε_i's are mutually independently distributed, $u_i \sim N(0, \sigma_u^2)$ and $\varepsilon_i \sim N(0, \sigma_i^2)$, respectively.

In matrix notation,

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u} + \boldsymbol{\varepsilon},$$

and $\boldsymbol{y} \sim N(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = \sigma_u^2 \boldsymbol{I}_m + \boldsymbol{D}$, $\boldsymbol{D} = diag(\sigma_1^2, \ldots, \sigma_m^2)$.

Multivariate Mixed Linear Model

Spatio-Temporal Small Area Model

The focus of this work is on the multivariate version of the extended Fay-Herriot model which includes spatial-temporal dependence structure. This extended model accommodates different patterns of spatial correlations and changes over time in order to improve estimation of the model parameters.

Spatio-Temporal Small Area Model

Let $\hat{\theta}_{it}$ be the direct estimator of the parameter of interest θ_{it}, $i = 1, \dots, m$ and $t = 1, \dots, T$, and the sampling model is the following:

$$\hat{\theta}_{it} = \theta_{it} + \varepsilon_{it},$$

where the vector $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \dots, \varepsilon_{iT})'$ of sampling errors for area i, $\boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \boldsymbol{\Psi}_i)$, where the covariance matrix $\boldsymbol{\Psi}_i$ is known. The linking model for the parameter of interest θ_{it} is

$$\theta_{it} = \boldsymbol{x}_{it}'\boldsymbol{\beta} + u_i + v_{it},$$

where u_i is a random area effect and v_{it} is an interaction area-by-time effect.

Multivariate Mixed Linear Model

Suppose that for a given unit (city, region) m distinct characteristics (small area means) are measured at each of t different occasions.

Multivariate Mixed Linear Model

Suppose that for a given unit (city, region) m distinct characteristics (small area means) are measured at each of t different occasions.

We assume that we have N units such that the measurements for different units are independent.

Multivariate Mixed Linear Model

Suppose that for a given unit (city, region) m distinct characteristics (small area means) are measured at each of t different occasions.

Let y_{ijk} denote the measurement of the ith characteristic at occasion j on unit k, $i = 1, \ldots, m$, $j = 1, \ldots, t$, $k = 1, \ldots, N$, and set $\boldsymbol{y}_{jk} = (y_{11k}, \ldots, y_{mjk})'$. Then we have the following model for $\boldsymbol{y}_{jk}$:

$$\boldsymbol{y}_{jk} = \boldsymbol{\theta}_k \boldsymbol{X}_j + \boldsymbol{u}_{jk},$$

or

$$\boldsymbol{Y}_k = \boldsymbol{\theta}_k \boldsymbol{X}' + \boldsymbol{U}_k,$$

where $\boldsymbol{X}' = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_t)$ is a known $q \times t$ matrix of full rank, $q \leq t$, and $\boldsymbol{\theta}_k$ is an $m \times q$ matrix.

Multivariate Mixed Linear Model

Using the vec operator, $\boldsymbol{y}_k = vec(Y_k)$, we can rewrite the model as following

$$\boldsymbol{y}_k = (\boldsymbol{X} \otimes \boldsymbol{I}_m)vec(\boldsymbol{\theta}_k) + \boldsymbol{u}_k,$$

where
$\boldsymbol{u}_k = vec(\boldsymbol{U}_k)$ and $vec(ABC) = (\boldsymbol{C}' \otimes \boldsymbol{A})vec(\boldsymbol{B})$.

Assuming that $(vec(\boldsymbol{\theta}_1), \ldots, vec(\boldsymbol{\theta}_N)) = \boldsymbol{B}\boldsymbol{A}'$, where $\boldsymbol{B} : mq \times r$ matrix of unknown parameters, and $\boldsymbol{A}' = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N)$ is an $r \times N$ matrix of known constants of full rank $r < N$.

Multivariate Mixed Linear Model

Let $\boldsymbol{Y}' = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N)$ and $\boldsymbol{U}' = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N)$, we get the following multivariate mixed linear model:

$$\boldsymbol{Y}' = (\boldsymbol{X} \otimes \boldsymbol{I}_m)\boldsymbol{B}\boldsymbol{A}' + \boldsymbol{U}'.$$

We assume that the columns of $\boldsymbol{U}'$ are independently distributed as $N(\mathbf{0}, \boldsymbol{\Omega})$, where $\boldsymbol{\Omega}$ is an unknown $mt \times mt$ covariance matrix.

Multivariate Mixed Linear Model

Let $\boldsymbol{Y}' = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N)$ and $\boldsymbol{U}' = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N)$, we get the following multivariate mixed linear model:

$$\boldsymbol{Y}' = (\boldsymbol{X} \otimes \boldsymbol{I}_m)\boldsymbol{B}\boldsymbol{A}' + \boldsymbol{U}'.$$

We assume that the columns of $\boldsymbol{U}'$ are independently distributed as $N(\boldsymbol{0}, \boldsymbol{\Omega})$, where $\boldsymbol{\Omega}$ is an unknown $mt \times mt$ covariance matrix.

If there are no special assumptions about the structure of the covariance matrix $\boldsymbol{\Omega}$, then we have the Growth Curve Model considered by Potthoff and Roy (1964), but involving multiple responses.

Multivariate Mixed Linear Model

- In many practical cases the dimension mt may be quite large relative to N.
- In this case a specific structure should be imposed on $\mathbf{\Omega}$ in order to obtain accurate estimates.
- In many cases a structured covariance matrix $\mathbf{\Omega}$ may be reasonable. Incorporating this covariance structure in the analysis would generally lead to more efficient inferences.

A structure of $\boldsymbol{\Omega}$ that may be appropriate to consider in some situations is a compound symmetry pattern. Under this structure, we have the following model

$$\boldsymbol{y}_{jk} = \boldsymbol{\theta}_k \boldsymbol{X}_j + \boldsymbol{\lambda}_k + \boldsymbol{\varepsilon}_{jk},$$

where $\boldsymbol{\lambda}_k$ is the $m \times 1$ vector of random effects associated with the kth unit, $\boldsymbol{\lambda}_k \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\lambda)$, independent of the random errors $\boldsymbol{\varepsilon}_{jk}$, $\boldsymbol{\varepsilon}_{jk} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_e)$. Observe that

$$\boldsymbol{Y}_k = \boldsymbol{\theta}_k \boldsymbol{X}' + \boldsymbol{\lambda}_k \mathbb{1}' + \boldsymbol{E}_k$$

or applying vec-operator

$$\boldsymbol{y}_k = (\boldsymbol{X} \otimes \boldsymbol{I}_m) vec(\boldsymbol{\theta}_k) + (\mathbb{1} \otimes \boldsymbol{I}_m) \boldsymbol{\lambda}_k + \boldsymbol{e}_k.$$

Here, $\boldsymbol{\Omega} = cov(\boldsymbol{y}_k) = (\mathbb{1}\mathbb{1}' \otimes \boldsymbol{\Sigma}_\lambda) + (\boldsymbol{I}_t \otimes \boldsymbol{\Sigma}_e)$.

Multivariate Mixed Linear Model

Now the full model may be expressed as

$$\boldsymbol{Y}' = (\boldsymbol{X} \otimes \boldsymbol{I}_m)\boldsymbol{B}\boldsymbol{A}' + (\mathbb{1} \otimes \boldsymbol{I}_m)\boldsymbol{\Lambda}' + \boldsymbol{E}',$$

where $\boldsymbol{\Lambda}' = (\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_N)$, $\boldsymbol{E}' = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N)$.
In vec-notation we have

$$\boldsymbol{y} = (\boldsymbol{A} \otimes \boldsymbol{X} \otimes \boldsymbol{I}_m)\boldsymbol{\beta} + \boldsymbol{u}$$

with $\boldsymbol{y} = vec(\boldsymbol{Y}')$, $\boldsymbol{u} = vec(\boldsymbol{U}')$, $\boldsymbol{\beta} = vec(\boldsymbol{B})$. for this model we have

$$cov(\boldsymbol{u}) = (\boldsymbol{I}_N \otimes \boldsymbol{\Omega}) = (\boldsymbol{I}_N \otimes \mathbb{1}\mathbb{1}' \otimes \boldsymbol{\Sigma}_\lambda) + (\boldsymbol{I}_N \otimes \boldsymbol{I}_t \otimes \boldsymbol{\Sigma}_e).$$

Multivariate Mixed Linear Model

For this model, the generalized least squares estimator of $\boldsymbol{\beta}$ is the same as the least squares estimator. The MLE of $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}} = ((\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}' \otimes (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \otimes \boldsymbol{I}_m)\boldsymbol{y},$$

$$\begin{aligned} \text{cov}(\hat{\boldsymbol{\beta}}) &= (\boldsymbol{A}'\boldsymbol{A})^{-1} \otimes (\boldsymbol{X}'\boldsymbol{X})^{-1} \otimes \boldsymbol{\Sigma}_e \\ &\quad + (\boldsymbol{A}'\boldsymbol{A})^{-1} \otimes (\boldsymbol{v}_1\boldsymbol{v}_1') \otimes \boldsymbol{\Sigma}_\lambda, \end{aligned}$$

where $\boldsymbol{v}_1 = (10\ldots0)$, and the MLE of $\boldsymbol{B}$ is given by

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

Multivariate Mixed Linear Model

Partitioning $\boldsymbol{B}$ as $(\boldsymbol{\mu}' : \boldsymbol{\Gamma}')'$, where $\boldsymbol{\mu} : m \times r$ and $\boldsymbol{\Gamma} : m(q-1) \times r$, we get

$$\hat{\boldsymbol{\mu}} = \frac{1}{t}(\mathbb{1}' \otimes \boldsymbol{I}_m)\boldsymbol{Y}'\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1} = \bar{\boldsymbol{Y}}'\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1},$$

$$\bar{\boldsymbol{Y}}' = (\bar{\boldsymbol{y}}_{.1}, \ldots, \bar{\boldsymbol{y}}_{.N}),\ \bar{\boldsymbol{y}}_{.k} = \frac{1}{t}\sum_{j=1}^{t} \boldsymbol{y}_{jk},$$

and

$$\hat{\boldsymbol{\Gamma}} = ((\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}' \otimes \boldsymbol{I}_m)\boldsymbol{Y}'\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}, \quad \boldsymbol{X} = (\boldsymbol{1} : \boldsymbol{Z}).$$

Multivariate Mixed Linear Model

Unbiased estimators of the covariance matrices:

$$\hat{\boldsymbol{\Sigma}}_e = \boldsymbol{S}_e/(N(t-1)-r(q-1)),$$
$$\hat{\boldsymbol{\Sigma}}_\lambda = \frac{1}{t}(\boldsymbol{S}_\lambda/(N-r)-\hat{\boldsymbol{\Sigma}}_e),$$

where

$$\boldsymbol{S}_e = \sum_{k=1}^{N}\sum_{j=1}^{t}(\boldsymbol{y}_{jk}-\bar{\boldsymbol{y}}_{.k}-(\boldsymbol{Z}_j'\otimes \boldsymbol{I}_m)\hat{\boldsymbol{\Gamma}}\boldsymbol{a}_k) \times(\boldsymbol{y}_{jk}-\bar{\boldsymbol{y}}_{.k}-(\boldsymbol{Z}_j'\otimes \boldsymbol{I}_m)\hat{\boldsymbol{\Gamma}}\boldsymbol{a}_k)'$$

and

$$\boldsymbol{S}_\lambda = t\sum_{k=1}^{N}(\bar{\boldsymbol{y}}_{.k}-\hat{\boldsymbol{\mu}}\boldsymbol{a}_k)(\bar{\boldsymbol{y}}_{.k}-\hat{\boldsymbol{\mu}}\boldsymbol{a}_k)'.$$

Generalization

Our model can be extended to a following general random effects model

$$\boldsymbol{Y}_k = \boldsymbol{\theta}_k \boldsymbol{X}' + \boldsymbol{\Lambda}_k \boldsymbol{X}' + \boldsymbol{\Xi}_k \boldsymbol{W}' + \boldsymbol{E}_k,$$

where $\boldsymbol{W}'$ is an $s \times t$ matrix of known constants of full rank s, with $q + s \leq t$, such that $\boldsymbol{X}'\boldsymbol{W} = 0$, and $\boldsymbol{\Lambda}_k : m \times q$ and $\boldsymbol{\Xi}_k : m \times s$ are matrices of random effects; $\boldsymbol{\Lambda}_k$, $\boldsymbol{\Xi}_k$, and $\boldsymbol{E}_k$ are mutually independent.

Alternative Spatio-Temporal Models

Our model can be extended to a following general random effects model

$$\boldsymbol{Y}_k = \boldsymbol{\theta}_k \boldsymbol{X}' + \boldsymbol{\Lambda}_k \boldsymbol{X}' + \boldsymbol{\Xi}_k \boldsymbol{W}' + \boldsymbol{E}_k,$$

where $\boldsymbol{W}'$ is an $s \times t$ matrix of known constants of full rank s, with $q + s \leq t$, such that $\boldsymbol{X}'\boldsymbol{W} = 0$, and $\boldsymbol{\Lambda}_k : m \times q$ and $\boldsymbol{\Xi}_k : m \times s$ are matrices of random effects; $\boldsymbol{\Lambda}_k$, $\boldsymbol{\Xi}_k$, and $\boldsymbol{E}_k$ are mutually independent.

The unbiased estimates of the covariance matrices can be obtained from the following estimating equations

$$\begin{aligned}
(N(t-q-s))^{-1}\boldsymbol{S}_e &= \hat{\boldsymbol{\Sigma}}_e, \\
(N-r)^{-1}\boldsymbol{S}_\lambda &= (\boldsymbol{X}'\boldsymbol{X})^{-1} \otimes \hat{\boldsymbol{\Sigma}}_e + \hat{\boldsymbol{\Sigma}}_\lambda, \\
N^{-1}\boldsymbol{S}_\Xi &= (\boldsymbol{W}'\boldsymbol{W})^{-1} \otimes \hat{\boldsymbol{\Sigma}}_e + \hat{\boldsymbol{\Sigma}}_\lambda
\end{aligned}$$