International School of Economics at TSU
Econometrics II
Lasha Chochua

Problem Set 1

Instructions: You are encouraged to solve the problems before the recitation. Additionally, you are encouraged to work in groups. It is **not mandatory** to submit solutions unless stated otherwise. However, if you would like to share your solution, I would be happy to review it.

Problem 1: Consider the stacked system of linear equations written in matrix notation as:

$$\mathbf{Y} = \overline{\mathbf{X}}\beta + \mathbf{e},$$

where:

- $\mathbf{Y} \in \mathbb{R}^{nm\times 1}$ is the stacked outcome vector,
- $\overline{\mathbf{X}} \in \mathbb{R}^{nm\times \bar{k}}$ is the block-diagonal regressor matrix,
- $\beta \in \mathbb{R}^{\bar{k}\times 1}$ is the parameter vector,
- $\mathbf{e} \in \mathbb{R}^{nm\times 1}$ is the stacked error vector.

Assume the following:

- $\mathbb{E}[\mathbf{e} \mid X] = 0$
- $\mathbb{E}[\mathbf{e}\mathbf{e}'] = \Omega = I_n \otimes \Sigma$, with $\Sigma \in \mathbb{R}^{m\times m}$ a symmetric positive definite matrix.

Derive the Generalized Least Squares (GLS) estimator for β, and show that it can be written as:

$$\hat{\beta}_{\text{gls}} = \left(\overline{\mathbf{X}}'(I_n \otimes \Sigma^{-1})\overline{\mathbf{X}}\right)^{-1}\left(\overline{\mathbf{X}}'(I_n \otimes \Sigma^{-1})\mathbf{Y}\right).$$

Solution

The model is:

$$\mathbf{Y} = \overline{\mathbf{X}}\beta + \mathbf{e}, \qquad \text{with } \mathbb{E}[\mathbf{e}\mathbf{e}'] = I_n \otimes \Sigma.$$

To apply GLS, we transform the model so that the transformed error has identity covariance matrix. Pre-multiply both sides by

$$(I_n \otimes \Sigma^{-1/2}),$$

which gives

$$(I_n \otimes \Sigma^{-1/2})\mathbf{Y} = (I_n \otimes \Sigma^{-1/2})\overline{\mathbf{X}}\beta + (I_n \otimes \Sigma^{-1/2})\mathbf{e}.$$

Define

$$\tilde{\mathbf{Y}} = (I_n \otimes \Sigma^{-1/2})\mathbf{Y},$$

and

$$\tilde{\mathbf{X}} = (I_n \otimes \Sigma^{-1/2})\overline{\mathbf{X}}.$$

Then the transformed model becomes

$$\tilde{\mathbf{Y}} = \tilde{\mathbf{X}}\beta + \tilde{\mathbf{e}},$$

where

$$\tilde{\mathbf{e}} = (I_n \otimes \Sigma^{-1/2})\mathbf{e}.$$

Its covariance matrix is

$$\mathbb{E}[\tilde{\mathbf{e}}\tilde{\mathbf{e}}'] = (I_n \otimes \Sigma^{-1/2})\mathbb{E}[\mathbf{e}\mathbf{e}'](I_n \otimes \Sigma^{-1/2})'.$$

Using $\mathbb{E}[\mathbf{e}\mathbf{e}'] = I_n \otimes \Sigma$, we obtain

$$\mathbb{E}[\tilde{\mathbf{e}}\tilde{\mathbf{e}}'] = (I_n \otimes \Sigma^{-1/2})(I_n \otimes \Sigma)(I_n \otimes \Sigma^{-1/2}) = I_n \otimes I_m = I_{nm}.$$

So the transformed model has spherical errors, and we can apply OLS:

$$\hat{\beta}_{\text{gls}} = (\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{Y}}.$$

Now substitute the definitions of $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$:

$$\hat{\beta}_{\text{gls}} = \left[\overline{\mathbf{X}}'(I_n \otimes \Sigma^{-1/2})'(I_n \otimes \Sigma^{-1/2})\overline{\mathbf{X}}\right]^{-1}\overline{\mathbf{X}}'(I_n \otimes \Sigma^{-1/2})'(I_n \otimes \Sigma^{-1/2})\mathbf{Y}.$$

Since Σ is symmetric positive definite, $\Sigma^{-1/2}$ is symmetric, so

$$(I_n \otimes \Sigma^{-1/2})' = I_n \otimes \Sigma^{-1/2}.$$

Hence,

$$(I_n \otimes \Sigma^{-1/2})'(I_n \otimes \Sigma^{-1/2}) = I_n \otimes \Sigma^{-1}.$$

Therefore,

$$\hat{\beta}_{\text{gls}} = \left(\overline{\mathbf{X}}'(I_n \otimes \Sigma^{-1})\overline{\mathbf{X}}\right)^{-1} \left(\overline{\mathbf{X}}'(I_n \otimes \Sigma^{-1})\mathbf{Y}\right).$$

This is the GLS estimator in matrix form.

Problem 2: Consider the stacked system

$$\mathbf{Y} = \overline{\mathbf{X}}\beta + \mathbf{e},$$

where

- $\mathbf{Y}$ is the $nm \times 1$ stacked vector of dependent variables,
- $\overline{\mathbf{X}}$ is the $nm \times \bar{k}$ stacked regressor matrix,
- $\mathbf{e}$ is the $nm \times 1$ stacked error vector.

Assume

$$E[\mathbf{e} \mid X] = 0$$

and

$$E[\mathbf{e}\mathbf{e}' \mid X] = I_n \otimes \Sigma,$$

where Σ is an $m \times m$ positive definite matrix.

The GLS estimator is

$$\hat{\beta}_{gls} = \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i\right)^{-1} \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} Y_i\right).$$

Show that

$$\mathrm{Var}(\widehat{\beta}_{gls} \mid X) = \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i \right)^{-1}.$$

Solution

Start from the GLS estimator

$$\widehat{\beta}_{gls} = \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i \right)^{-1} \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} Y_i \right).$$

Substitute the model

$$Y_i = \overline{X}_i \beta + e_i.$$

Then

$$\widehat{\beta}_{gls} = \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i \right)^{-1} \sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} (\overline{X}_i \beta + e_i).$$

Expanding,

$$\widehat{\beta}_{gls} = \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i \right)^{-1} \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i \right) \beta + \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i \right)^{-1} \sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} e_i.$$

Hence

$$\widehat{\beta}_{gls} - \beta = \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i \right)^{-1} \sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} e_i.$$

Now take the conditional variance given X:

$$\mathrm{Var}(\widehat{\beta}_{gls} \mid X) = \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i \right)^{-1} \mathrm{Var}\left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} e_i \,\middle|\, X \right) \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i \right)^{-1}.$$

Because observations are independent across i and

$$\operatorname{Var}(e_i \mid X) = \Sigma,$$

we have

$$\operatorname{Var}\left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} e_i \,\middle|\, X\right) = \sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \operatorname{Var}(e_i \mid X) \Sigma^{-1} \overline{X}_i = \sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \Sigma \Sigma^{-1} \overline{X}_i.$$

Therefore

$$\operatorname{Var}\left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} e_i \,\middle|\, X\right) = \sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i.$$

Substituting back,

$$\operatorname{Var}(\hat{\beta}_{gls} \mid X) = \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i\right)^{-1} \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i\right) \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i\right)^{-1}.$$

Hence,

$$\operatorname{Var}(\hat{\beta}_{gls} \mid X) = \left(\sum_{i=1}^{n} \overline{X}_i' \Sigma^{-1} \overline{X}_i\right)^{-1}.$$

Problem 3: Suppose the covariance matrix Σ in the system model is unknown. We consider the same system as in Problem 1:

$$\mathbf{Y} = \overline{\mathbf{X}}\beta + \mathbf{e},$$

We still assume:

- $\mathbb{E}[\mathbf{e} \mid X] = 0$
- $\mathbb{E}[\mathbf{e}\mathbf{e}' \mid X] = \Omega = I_n \otimes \Sigma$
- Σ is **unknown**, but we have a **consistent estimator** $\hat{\Sigma}$

1. Write the **feasible GLS (FGLS)** estimator, $\hat{\beta}_{\text{sur}}$, using matrix notation.
2. Prove that under regularity conditions (e.g. Assumption 7.2 and consistency of $\hat{\Sigma}$), the FGLS estimator is **asymptotically normal**:

$$\sqrt{n}(\hat{\beta}_{\text{sur}} - \beta) \xrightarrow{d} \mathcal{N}(0, V_\beta^*)$$

where

$$V_\beta^* = \left(\operatorname{plim} \frac{1}{n} \overline{\mathbf{X}}' (I_n \otimes \Sigma^{-1}) \overline{\mathbf{X}}\right)^{-1}.$$

Solution

1. Feasible GLS Estimator

The **feasible GLS (FGLS)** or **SUR** estimator is given by

$$\hat{\beta}_{\text{sur}} = \left(\overline{\mathbf{X}}' (I_n \otimes \hat{\Sigma}^{-1}) \overline{\mathbf{X}}\right)^{-1} \overline{\mathbf{X}}' (I_n \otimes \hat{\Sigma}^{-1}) \mathbf{Y}.$$

Here, $\hat{\Sigma}$ is a consistent estimator of Σ, typically constructed using OLS residuals from each equation:

$$\hat{u}_i = Y_i - \overline{X}_i \hat{\beta}^{\text{OLS}},$$

and

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^{n} \hat{u}_i \hat{u}_i'.$$

2. Asymptotic Distribution

Start from the FGLS estimator:

$$\hat{\beta}_{\text{sur}} = \left(\overline{\mathbf{X}}' (I_n \otimes \hat{\Sigma}^{-1}) \overline{\mathbf{X}}\right)^{-1} \overline{\mathbf{X}}' (I_n \otimes \hat{\Sigma}^{-1}) \mathbf{Y}.$$

Substitute the model

$$\mathbf{Y} = \overline{\mathbf{X}} \beta + \mathbf{e}.$$

Then

$$\hat{\beta}_{\text{sur}} = \left(\overline{\mathbf{X}}'(I_n \otimes \hat{\Sigma}^{-1})\overline{\mathbf{X}}\right)^{-1} \overline{\mathbf{X}}'(I_n \otimes \hat{\Sigma}^{-1})(\overline{\mathbf{X}}\beta + \mathbf{e}).$$

Hence,

$$\hat{\beta}_{\text{sur}} - \beta = \left(\overline{\mathbf{X}}'(I_n \otimes \hat{\Sigma}^{-1})\overline{\mathbf{X}}\right)^{-1} \overline{\mathbf{X}}'(I_n \otimes \hat{\Sigma}^{-1})\mathbf{e}.$$

Multiply both sides by $\sqrt{n}$:

$$\sqrt{n}(\hat{\beta}_{\text{sur}} - \beta) = \left(\frac{1}{n}\overline{\mathbf{X}}'(I_n \otimes \hat{\Sigma}^{-1})\overline{\mathbf{X}}\right)^{-1} \left(\frac{1}{\sqrt{n}}\overline{\mathbf{X}}'(I_n \otimes \hat{\Sigma}^{-1})\mathbf{e}\right).$$

Now analyze the two factors separately.

Because $\hat{\Sigma} \xrightarrow{p} \Sigma$, we have

$$\hat{\Sigma}^{-1} \xrightarrow{p} \Sigma^{-1}.$$

Therefore,

$$\frac{1}{n}\overline{\mathbf{X}}'(I_n \otimes \hat{\Sigma}^{-1})\overline{\mathbf{X}} - \frac{1}{n}\overline{\mathbf{X}}'(I_n \otimes \Sigma^{-1})\overline{\mathbf{X}} \xrightarrow{p} 0.$$

Under the law of large numbers,

$$\frac{1}{n}\overline{\mathbf{X}}'(I_n \otimes \Sigma^{-1})\overline{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n} \overline{X}_i'\Sigma^{-1}\overline{X}_i \xrightarrow{p} Q,$$

where

$$Q = \operatorname{plim} \frac{1}{n}\overline{\mathbf{X}}'(I_n \otimes \Sigma^{-1})\overline{\mathbf{X}}.$$

Thus,

$$\left(\frac{1}{n}\overline{\mathbf{X}}'(I_n \otimes \hat{\Sigma}^{-1})\overline{\mathbf{X}}\right)^{-1} \xrightarrow{p} Q^{-1}.$$

Next consider

$$\frac{1}{\sqrt{n}}\overline{\mathbf{X}}'(I_n \otimes \hat{\Sigma}^{-1})\mathbf{e}.$$

By consistency of $\hat{\Sigma}^{-1}$, replacing $\hat{\Sigma}^{-1}$ with Σ^{-1} only creates an $o_p(1)$ difference, so

$$\frac{1}{\sqrt{n}}\overline{\mathbf{X}}'(I_n \otimes \hat{\Sigma}^{-1})\mathbf{e} = \frac{1}{\sqrt{n}}\overline{\mathbf{X}}'(I_n \otimes \Sigma^{-1})\mathbf{e} + o_p(1).$$

Using the block structure,

$$\frac{1}{\sqrt{n}}\overline{\mathbf{X}}'(I_n \otimes \Sigma^{-1})\mathbf{e} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\overline{X}_i'\Sigma^{-1}e_i.$$

By the central limit theorem,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\overline{X}_i'\Sigma^{-1}e_i \xrightarrow{d} \mathcal{N}(0, \Psi),$$

where

$$\Psi = \mathbb{E}\left[\overline{X}_i'\Sigma^{-1}e_ie_i'\Sigma^{-1}\overline{X}_i\right].$$

Since

$$\mathbb{E}[e_ie_i' \mid X] = \Sigma,$$

it follows that

$$\Psi = \mathbb{E}\left[\overline{X}_i'\Sigma^{-1}\Sigma\Sigma^{-1}\overline{X}_i\right] = \mathbb{E}\left[\overline{X}_i'\Sigma^{-1}\overline{X}_i\right].$$

Equivalently, in stacked notation this is exactly the probability limit

$$Q = \operatorname{plim}\frac{1}{n}\overline{\mathbf{X}}'(I_n \otimes \Sigma^{-1})\overline{\mathbf{X}}.$$

Therefore,

$$\frac{1}{\sqrt{n}}\overline{\mathbf{X}}'(I_n \otimes \hat{\Sigma}^{-1})\mathbf{e} \xrightarrow{d} \mathcal{N}(0, Q).$$

Combining the two pieces and applying Slutsky's theorem,

$$\sqrt{n}(\hat{\beta}_{\text{sur}} - \beta) \xrightarrow{d} \mathcal{N}(0, Q^{-1}QQ^{-1}) = \mathcal{N}(0, Q^{-1}).$$

Hence,

$$\sqrt{n}(\hat{\beta}_{\text{sur}} - \beta) \xrightarrow{d} \mathcal{N}(0, V_{\beta}^{*}),$$

where

$$V_{\beta}^{*} = \left(\operatorname{plim} \frac{1}{n} \overline{\mathbf{X}}'(I_n \otimes \Sigma^{-1})\overline{\mathbf{X}}\right)^{-1}.$$

Problem 4: Use the data in NBASAL.RAW to answer this question.

a. Estimate an SUR model for the three response variables points, rebounds, and assists. The explanatory variables in each equation should be age, exper, exper2, guard, forward, black, and marr. Does marital status have a positive or negative a¤ect on each variable? Is it statistically significant in the assists equation?

b. Test the hypothesis that marital status can be excluded entirely from the system.

Note: For answer see the Jupyter Notebook.