Seminar 2

Seemingly Unrelated Regression (SUR)

Petr Gapko

(petr.gapko@seznam.cz)

Linear Algebra recap : inverse matrix

i) The inverse of a 2x2 matrix:

Original matrix: $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ Inverse matrix: $A^{-1} = \frac{1}{|A|} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$, where $|A| = a_{11}\, a_{22} - a_{12}\, a_{21}$

ii) The inverse of a partitioned matrix:

Original matrix: $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ Inverse matrix: $A^{-1} = \begin{pmatrix} E & -E\, A_{12}\, A_{22}^{-1} \\ -A_{22}^{-1}\, A_{21}\, E & A_{22}^{-1} + A_{22}^{-1}\, A_{21}\, E\, A_{12}\, A_{22}^{-1} \end{pmatrix}$,

where $E = \left(A_{11} - A_{12}\, A_{22}^{-1}\, A_{21}\right)^{-1}$. Alternatively it can be expressed as $A^{-1} = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}\, A_{12}\, F\, A_{21}\, A_{11}^{-1} & -A_{11}^{-1}\, A_{12}\, F \\ -F\, A_{21}\, A_{11}^{-1} & F \end{pmatrix}$, where $F = \left(A_{22} - A_{21}\, A_{11}^{-1}\, A_{12}\right)^{-1}$.

Neccessary and Sufficient Condition for OLS to be Equivalent to GLS

Derivation from Milliken and Albohali (1984): Let's have a regression model y=βX+ϵ . From this model we derive $\hat{\beta}_{\text{GLS}} = \left(X^T \Omega^{-1} X\right)^{-1} X^T \Omega^{-1}$y. We can rewrite $y = P_x\, y + \overline{P}_x\, y$, where $\overline{P}_x$ is an orthogonal projection matrix of X and is equal to $I_n - P_x$ and $P_x = X\left(X^T X\right)^{-1} X^T$. By replacing y we get

$$\hat{\beta}_{\text{GLS}} = \left(X^T \Omega^{-1} X\right)^{-1} X^T \Omega^{-1}\left[P_x\, y + \overline{P}_x\, y\right] = \hat{\beta}_{\text{OLS}} + \left(X^T \Omega^{-1} X\right)^{-1} X^T \Omega^{-1}\, \overline{P}_x\, y$$

The last term can be zero (and thus $\hat{\beta}_{\text{GLS}} = \hat{\beta}_{\text{OLS}}$) for every y if and only if $X^T \Omega^{-1}\, \overline{P}_x = 0$.

Frish - Waugh - Lovell Theorem

This Theorem says:

(i) that the determination of the coefficients in a standard regression model via ordinary least squares and a method involving projection matrices are equivalent

(ii) the residuals are identical too

Problem #1 from Baltagi, Chapter 10

(a) Show that the OLS on a system of two Zellner's SUR equations:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

is the same as OLS on each equation taken separately. What about estimated variance-covariance matrix of the coeffs? Will they be the same?

Solution:

OLS on each regression separately:

eq.1: $\hat{\beta}_{1,\text{OLS}} = \left(X_1{}^T X_1\right)^{-1} X_1{}^T y_1$

eq.2: $\hat{\beta}_{2,\text{OLS}} = \left(X_2{}^T X_2\right)^{-1} X_2{}^T y_2$

OLS on the system of two Zelner's SUR equations:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

$$\hat{\beta}_{\text{OLS}} = \left(\begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix}^T \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \right)^{-1} \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix}^T \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \left(\begin{pmatrix} X_1^T & 0 \\ 0 & X_2^T \end{pmatrix} \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \right)^{-1} \begin{pmatrix} X_1^T & 0 \\ 0 & X_2^T \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} =$$

$$\begin{pmatrix} X_1^T X_1 & 0 \\ 0 & X_2^T X_2 \end{pmatrix}^{-1} \begin{pmatrix} X_1^T y_1 \\ X_2^T y_2 \end{pmatrix} = \begin{pmatrix} \left(X_1^T X_1\right)^{-1} & 0 \\ 0 & \left(X_2^T X_2\right)^{-1} \end{pmatrix} \begin{pmatrix} X_1^T y_1 \\ X_2^T y_2 \end{pmatrix} = \begin{pmatrix} \left(X_1^T X_1\right)^{-1} X_1^T y_1 \\ \left(X_2^T X_2\right)^{-1} X_2^T y_2 \end{pmatrix}$$

What about the var-cov matrix?

Separate equations:

$e_{1,\text{OLS}} = y_1 - X_1 \hat{\beta}_{1,\text{ OLS}}$. s_1 is an estimate of σ_1 and we can write that: $\text{var}(\hat{\beta}_{1,\text{OLS}}) = s_1^2\left(X_1^T X_1\right)^{-1}$, $s_1^2 = e_{1,\text{OLS}}^T e_{1,\text{OLS}} / (T - K_1)$

eq. 2 is similar: $\text{var}(\hat{\beta}_{1,\text{OLS}}) = s_2^2\left(X_2^T X_2\right)^{-1}$.

System of equations:

$\text{RSS} = e_{1,\text{OLS}}^T e_{1,\text{OLS}} + e_{2,\text{OLS}}^T e_{2,\text{OLS}}$

$s^2 = \frac{1}{2\,T - K_1 - K_2} \left(e_{1,\text{OLS}}^T e_{1,\text{OLS}} + e_{2,\text{OLS}}^T e_{2,\text{OLS}}\right)$

$\hat{\text{var}}\ \hat{\beta} = s^2 \begin{pmatrix} \left(X_1^T X_1\right)^{-1} & 0 \\ 0 & \left(X_2^T X_2\right)^{-1} \end{pmatrix}$.

Conclusion: estimated var-cov matrices differ by constant s^2instead of s_1 or s_2.

(b) In the General Linear Model, we found a necessary and sufficient condition for OLS to be equivalent to GLS. Show that a necessary and sufficient condition for Zellner's GLS to be equivalent to OLS is that $\sigma^{\text{ij}} X_i^T \overline{P}_{x_j} = 0$ for $i \neq j$.

Solution: The case of SUR - the case of 2 eq's:

$$X = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix}, \quad \overline{P_X} = I_{2\,T} - P_X$$

$$X^T \Omega^{-1} \overline{P}_x = 0, \quad \Omega^{-1} = \Sigma^{-1} \otimes I_T, \text{ where } \Sigma^{-1} = \begin{pmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{21} & \sigma^{22} \end{pmatrix}$$

$$\begin{pmatrix} X_1^T & 0 \\ 0 & X_2^T \end{pmatrix} \begin{pmatrix} \sigma^{11}\, I & \sigma^{12}\, I \\ \sigma^{21}\, I & \sigma^{22}\, I \end{pmatrix} \begin{pmatrix} \overline{P_{X_1}} & 0 \\ 0 & \overline{P_{X_2}} \end{pmatrix} = \begin{pmatrix} \sigma^{11}\, X_1^T\, \overline{P_{X_1}} & \sigma^{12}\, X_1^T\, \overline{P_{X_2}} \\ \sigma^{21}\, X_2^T\, \overline{P_{X_1}} & \sigma^{22}\, X_2^T\, \overline{P_{X_2}} \end{pmatrix}$$

If we denote matrix $A = X^T\, \Omega^{-1}\, \overline{P}_x$, the element A_{ij} can be expressed as $\sigma^{\text{ij}}\, X_i^T\, \overline{P_{X_j}}$. The condition we have can be rewritten as $\sigma^{\text{ij}}\, X_i^T\, \overline{P_{X_j}} = A_{\text{ij}} = 0$. If $i = j$ it is obvious that this condition holds. If $i \neq j$ we have two cases: $\sigma^{\text{ij}}\, X_i^T\, \overline{P}_{x_j} = 0$ if and only if:

i) $\sigma^{\text{ij}} = 0$ or

ii) $X_i^T\, \overline{P}_{x_j} = 0$, i.e. $X_i = X_j$

(c) Show that the two sufficient conditions given by Zellner for SUR to be equivalent to OLS both satisfy the necessary and sufficient condition given in part (b)

Solution:

Σ is diagonal $\Rightarrow \Sigma^{-1}$ is diagonal (in other words: $\sigma^{ij} = 0$) corresponds to i)

if $X_i = X_j$ for all i,j then ii) is satisfied

(d) Show that if $X_i = X_j C^T$ where C is an arbitrary nonsingular matrix, then the necessary and sufficient condition given in part (b) is satisfied.

Solution:

$$P_{X_i} = X_i\left(X_i^T X_i\right)^{-1} X_i^T = X_j C^T\left(\left(X_j C^T\right)^T X_j C^T\right)^{-1} \left(X_j C^T\right)^T =$$
$$X_j C^T\left(C X_j^T X_j C^T\right)^{-1} C X_j^T = X_j C^T\left(C^T\right)^{-1} \left(X_j^T X_j\right)^{-1} C^{-1} C X_j^T = X_j\left(X_j^T X_j\right)^{-1} X_j^T = P_{X_j} \Rightarrow \overline{P_{X_i}} = \overline{P_{X_j}}$$

$$\sigma^{\text{ij}} X_i^T \overline{P_{X_j}} = \sigma^{\text{ij}}\left(X_j C^T\right)^T \overline{P_{X_j}} = \sigma^{\text{ij}} C X_j^T \overline{P_{X_j}} = 0$$

Problem #5 from Baltagi, Chapter 10

(a) Show that var($\hat{\beta}_{12,\text{OLS}}$) = $\frac{\sigma_{11}}{m_{x_1 x_1}}$ and var($\hat{\beta}_{22,\text{OLS}}$) = $\frac{\sigma_{22}}{m_{x_2 x_2}}$, where $m_{x_i x_j} = \sum_{t=1}^{T} \left(X_{\text{it}} - \overline{X_i}\right)\left(X_{\text{jt}} - \overline{X_j}\right)$ for $i,\ j = 1,\ 2$.

Solution: $\hat{\beta}_{12,\text{OLS}} = \left(X_1^T A X_1\right)^{-1} X_1^T A Y_1,\ A = \left(I_T - \frac{i\, i^T}{T}\right),\ X_1^T A X_1 = \sum_{t=1}^{T} \left(X_{1\,t} - \overline{X_1}\right)\left(X_{1\,t} - \overline{X_1}\right) = m_{x_1 x_1}$

var($\hat{\beta}_{12,\text{OLS}}$) $= \sigma_{11}\left(X_1^T A X_1\right)^{-1} = \frac{\sigma_{11}}{m_{x_1 x_1}}$. Similarly we proof that var($\hat{\beta}_{22,\text{OLS}}$) = $\frac{\sigma_{22}}{m_{x_2 x_2}}$

(b) $\text{var}\begin{pmatrix}\hat{\beta}_{12,\text{GLS}}\\ \hat{\beta}_{22,\text{GLS}}\end{pmatrix} = (\sigma_{11}\,\sigma_{22} - \sigma_{12}^2)\begin{pmatrix}\sigma_{22}\,m_{x_1\,x_1} & -\sigma_{12}\,m_{x_1\,x_2}\\ -\sigma_{12}\,m_{x_1\,x_2} & \sigma_{11}\,m_{x_2\,x_2}\end{pmatrix}^{-1}$. Deduce that $\text{var}(\hat{\beta}_{12,\text{GLS}}) = (\sigma_{11}\,\sigma_{22} - \sigma_{12}^2)\,\sigma_{11}\,m_{x_2\,x_2} / (\sigma_{11}\,\sigma_{22}\,m_{x_2\,x_2}\,m_{x_1\,x_1} - \sigma_{12}^2\,m_{x_1\,x_2}^2)$ and $\text{var}(\hat{\beta}_{22,\text{GLS}}) = (\sigma_{11}\,\sigma_{22} - \sigma_{12}^2)\,\sigma_{22}\,m_{x_1\,x_1} / (\sigma_{11}\,\sigma_{22}\,m_{x_1\,x_1}\,m_{x_2\,x_2} - \sigma_{12}^2\,m_{x_1\,x_2}^2)$

Solution:

$$\begin{pmatrix}\hat{\beta}_{12,\text{GLS}}\\ \hat{\beta}_{22,\text{GLS}}\end{pmatrix} =$$

$$\left(\begin{pmatrix}X_1^T A & 0\\ 0 & X_2^T A\end{pmatrix}\begin{pmatrix}\sigma^{11}\otimes I & \sigma^{12}\otimes I\\ \sigma^{21}\otimes I & \sigma^{22}\otimes I\end{pmatrix}\begin{pmatrix}A X_1 & 0\\ 0 & A X_2\end{pmatrix}\right)^{-1}\begin{pmatrix}X_1^T A & 0\\ 0 & X_2^T A\end{pmatrix}\begin{pmatrix}\sigma^{11}\otimes I & \sigma^{12}\otimes I\\ \sigma^{21}\otimes I & \sigma^{22}\otimes I\end{pmatrix}\begin{pmatrix}Y_1\\ Y_2\end{pmatrix} =$$

$$\left(\begin{pmatrix}X_1^T A & 0\\ 0 & X_2^T A\end{pmatrix}\begin{pmatrix}\sigma^{11} & \sigma^{12}\\ \sigma^{21} & \sigma^{22}\end{pmatrix}\begin{pmatrix}A X_1 & 0\\ 0 & A X_2\end{pmatrix}\right)^{-1}\begin{pmatrix}X_1^T A & 0\\ 0 & X_2^T A\end{pmatrix}\begin{pmatrix}\sigma^{11} & \sigma^{12}\\ \sigma^{21} & \sigma^{22}\end{pmatrix}\begin{pmatrix}Y_1\\ Y_2\end{pmatrix} =$$

$$\begin{pmatrix}\sigma^{11}\,X_1^T\,A\,X_1 & \sigma^{12}\,X_1^T\,A\,X_2\\ \sigma^{21}\,X_2^T\,A\,X_1 & \sigma^{22}\,X_2^T\,A\,X_2\end{pmatrix}^{-1}\begin{pmatrix}\sigma^{11}\,X_1^T\,A\,Y_1 & \sigma^{12}\,X_1^T\,A\,Y_2\\ \sigma^{21}\,X_2^T\,A\,Y_1 & \sigma^{22}\,X_2^T\,A\,Y_2\end{pmatrix}$$

$$\text{var}\begin{pmatrix}\hat{\beta}_{12,\text{GLS}}\\ \hat{\beta}_{22,\text{GLS}}\end{pmatrix} = \begin{pmatrix}\sigma^{11}\,X_1^T\,A\,X_1 & \sigma^{12}\,X_1^T\,A\,X_2\\ \sigma^{21}\,X_2^T\,A\,X_1 & \sigma^{22}\,X_2^T\,A\,X_2\end{pmatrix}^{-1} = \begin{pmatrix}\sigma^{11}\,m_{x_1\,x_1} & \sigma^{12}\,m_{x_1\,x_2}\\ \sigma^{21}\,m_{x_2\,x_1} & \sigma^{22}\,m_{x_2\,x_2}\end{pmatrix} =$$

$$(\sigma_{11}\,\sigma_{22} - \sigma_{21}^2)\begin{pmatrix}\sigma_{22}\,m_{x_1\,x_1} & -\sigma_{12}\,m_{x_1\,x_2}\\ -\sigma_{21}\,m_{x_2\,x_1} & \sigma_{11}\,m_{x_2\,x_2}\end{pmatrix}^{-1} = \frac{\sigma_{11}\,\sigma_{22}-\sigma_{21}^2}{\sigma_{11}\,\sigma_{22}\,m_{x_1\,x_1}-\sigma_{21}^2\,m_{x_1\,x_2}^2}\begin{pmatrix}\sigma_{22}\,m_{x_1\,x_1} & \sigma_{12}\,m_{x_1\,x_2}\\ \sigma_{21}\,m_{x_2\,x_1} & \sigma_{11}\,m_{x_2\,x_2}\end{pmatrix}$$

$\text{var}(\hat{\beta}_{12,\text{GLS}}) = \frac{(\sigma_{11}\,\sigma_{22}-\sigma_{21}^2)\,\sigma_{11}\,m_{x_2\,x_2}}{\sigma_{11}\,\sigma_{22}\,m_{x_1\,x_1}\,m_{x_2\,x_2}-\sigma_{21}^2\,m_{x_1\,x_2}^2}$ and similarly $\text{var}(\hat{\beta}_{22,\text{GLS}}) = \frac{(\sigma_{11}\,\sigma_{22}-\sigma_{12}^2)\,\sigma_{22}\,m_{x_1\,x_1}}{(\sigma_{11}\,\sigma_{22}\,m_{x_1\,x_1}\,m_{x_2\,x_2}-\sigma_{12}^2\,m_{x_1\,x_2}^2)}$

(c) Using $\rho = \sigma_{12} / \left(\sigma_{11}\,\sigma_{22} - \sigma_{12}^2\right)^{1/2}$ and $r = m_{x_1 x_2} / (m_{x_1 x_1}\, m_{x_2 x_2})^{1/2}$ and the results in part (a) and (b), show that $\mathrm{var}(\hat{\beta}_{12,\mathrm{GLS}}) / \mathrm{var}(\hat{\beta}_{12,\mathrm{OLS}}) = \left(1 - \rho^2\right) / \left(1 - \rho^2\, r^2\right)$.

S o l u t i o n :

$\mathrm{var}(\hat{\beta}_{12,\mathrm{GLS}}) / \mathrm{var}(\hat{\beta}_{12,\mathrm{OLS}}) =$

$$\mathrm{var}(\hat{\beta}_{12,\mathrm{GLS}}) = \frac{\left(\sigma_{11}\,\sigma_{22} - \sigma_{21}^2\right)\sigma_{11}\, m_{x_2 x_2}}{\sigma_{11}\,\sigma_{22}\, m_{x_1 x_1}\, m_{x_2 x_2} - \sigma_{21}^2\, m_{x_1 x_2}^2} \Big/ \frac{\sigma_{22}}{m_{x_1 x_1}} = \frac{\left(\sigma_{11}\,\sigma_{22} - \sigma_{21}^2\right) m_{x_1 x_1}\, m_{x_2 x_2}}{\sigma_{11}\,\sigma_{22}\, m_{x_1 x_1}\, m_{x_2 x_2} - \sigma_{21}^2\, m_{x_1 x_2}^2} = \frac{\sigma_{12}^2 \left(\frac{1}{\rho^2} - 1\right) \frac{m_{x_1 x_2}^2}{r^2}}{\sigma_{12}^2\, \frac{1}{\rho^2}\, \frac{m_{x_1 x_2}^2}{r^2} - \sigma_{12}^2\, m_{x_1 x_2}^2} = \frac{\frac{1}{\rho^2 r^2} - \frac{1}{r^2}}{\frac{1}{\rho^2 r^2} - 1} = \frac{1-\rho^2}{1-\rho^2\, r^2}$$

(d) Differentiate $\text{var}(\hat{\beta}_{12,\text{GLS}}) / \text{var}(\hat{\beta}_{12,\text{OLS}}) = (1 - \rho^2)/(1 - \rho^2 r^2)$ with respect to $\theta = \rho^2$ and show that this expression is a non-increasing function of θ. Similarly, differenetiate the expression with respect to $\lambda = r^2$ and show that it is a non-decreasing function of λ. Finally, compute this efficiency measure for various values of ρ^2 and r^2 between 0 and 1 at 0,1 intervals.

Solution: $\partial\left(\frac{1-\rho^2}{1-\rho^2 r^2}\right) / \partial \rho^2 = \frac{-(1-\rho^2 r^2)+(1-\rho^2)(r^2)}{(1-\rho^2 r^2)^2} = \frac{r^2-1}{(1-\rho^2 r^2)^2}$. Both $\rho^2$and $r^2$are from interval $<0, 1>$ and thus the expression is non-positive for all values ρ^2 and r^2. Moreover the sign at ρ^2 is negative and this with the power of two causes that the expression is a non-increasing function in ρ^2.

$\partial\left(\frac{1-\rho^2}{1-\rho^2 r^2}\right) / \partial r^2 = \frac{\rho^2(1-\rho^2)}{(1-\rho^2 r^2)^2}$. As both $\rho^2$and $r^2$are from interval $<0, 1>$ this function will be non-negative for all values of ρ^2 and r^2 and thus the bigger r^2 the bigger value of the function.