14.383

Fall 2003.

9/13 - Handout #1

(thanks to:
Ziad Nejmeldeen).

TA: Alexis León
aleon@mit.edu
E52-351
(Off. hrs.: Mon., 5.30-7 pm)
in E52-242
- Freeman Room -

Section 1

1. Review of 14.382
   (a) Linear Algebra
   (b) Convergence Concepts
   (c) Laws of Large Numbers, Limit Theorems
2. Asymptotic Theory / Small Sample Theory
3. Brief Overview of Instrumental Variables (IV)

1(a) Linear Algebra

Consider $\underset{N\times 1}{Y} = \underset{N\times K}{X}\,\underset{K\times 1}{\beta} + \underset{N\times 1}{\epsilon}$ N observations, k regressors

i) The PROJECTION MATRIX $P_X = X(X'X)^{-1}X'$

$P_X Y$ projects Y onto the subspace spanned by X. This gives the predicted values $\hat{Y}$ in the sense of least squares.

ii) The ORTHOGONAL PROJECTION matrix $Q_X = I - P$

$Q_X Y$ projects Y onto the subspace orthogonal to X; $Q_X Y$ gives the part of Y that can't be explained by X. $Q_X X = 0$.

$Q_X Y = (I - P_X)Y = \hat{\epsilon}$, the least squares residual

iii) A matrix M is IDEMPOTENT if it is square and if $M^2 = M$. Both P_X and Q_X are idempotent.

iv) KRONECKER PRODUCTS

$$A: m \times n \qquad A \equiv [a_{ij}]$$
$$B: p \times q \qquad B = [b_{ij}]$$

$$A \otimes B = \underbrace{\begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}}_{mp \times nq}$$

- $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$, assuming the matrices are comformable.
- $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$, as long as A, B are square and inventible
- $(A \otimes B)' = A' \otimes B'$

v) RANK of a MATRIX

- The rank of a matrix is the number of linearly independent vectors (column vectors) in the matrix. For $\underset{n \times k}{A}$, rank$(A) \leq k$.
- If $\underset{n \times k}{A}$ has full rank and $n > k$, then $\underset{k \times k}{A'A}$ is psd and $\underset{n \times n}{AA'}$ is nonnegative definite.

1(b) Convergence Concepts

(i) Convergence in Probability

For a sequence of random variables $(Z_1, Z_2, \ldots, Z_n)$, if $\exists\, \delta > 0$ and a real number Z_0 such that

$$\lim_{n \to \infty} P\left(\| Z_n - Z_0 \| < \delta \right) = 1$$

(where $\| \cdot \|$ is the Euclidean norm), we say that $Z_n \xrightarrow{p} Z_0 \equiv \operatorname{plim} Z_n = Z_0$, or, Z_n CONVERGES in PROBABILITY to Z_0.

(ii) Convergence in Distribution

Z_n CONVERGES in DISTRIBUTION to Z_0 if $\lim_{n \to \infty} \left| P(Z_n \le z) - P(Z_0 \le z) \right| = 0 \quad \forall z$.

i.e. $Z_n \xrightarrow{d} Z_0$.

1(c) Laws of Large Numbers

The laws of Large Numbers specify the condition under which sample averages converge in probability to their population counter parts.

(i) For a random sample $Z_1, \ldots, Z_n$ such that $E(Z_n) = \mu < \infty$, the Weak Law of Large Numbers (WLLN) is that $\operatorname{plim}_{n \to \infty} \bar{Z}_n = \mu$.

(ii) Central Limit Theorems

These theorems specify the conditions under which sample averages are normally distributed as $n \to \infty$. These conditions are typically iid random variables, $E(|Z_i|^2) < \infty$, etc..

The CLT's are of the form:

$$\frac{\bar{Y} - E(\bar{Y})}{\sqrt{Var(\bar{Y})}} \xrightarrow{d} N(0,1).$$

2. Asymptotic and Small Sample Theory

(i) Small-Sample Results

$$\underset{n\times 1}{y} = \underset{n\times k}{X}\,\underset{k\times 1}{\beta} + \underset{n\times 1}{\epsilon}$$

(GM1) X has full column rank K

(GM2) $E(\varepsilon | X) = 0$

(GM3) $E(\varepsilon\varepsilon' | X) = \sigma^2 I$

Thm: Under GM1 - GM3, $\hat{\beta}_{OLS} = (X'X)^{-1}X'y$ is BLUE

(GM4) $\varepsilon \sim N(0, \sigma^2 I)$

Thm Under GM1 - GM4, $\hat{\beta}_{OLS}$ is BUE

(ii) Asymptotic Results

(OLS 1) $X'X/n \xrightarrow{p} M$ nonsingular

(OLS 2) $X'\epsilon/n \xrightarrow{p} 0$

(OLS 3) $X'\epsilon/\sqrt{n} \xrightarrow{d} N(0, V)$

(OLS 4) No Serial Correlation: $\forall\ t \neq s$

$$E(\varepsilon_t \varepsilon_s \mid X_t, X_s) = 0$$

(OLS 5) Homoskedasticity: $\forall\ t,\ E(\varepsilon_t^2 \mid X_t) = \sigma^2$

Thm (Consistency): Under OLS 1 – OLS 2, $\hat{\beta}_{OLS} \xrightarrow{p} \beta$

Thm (Asym. Normality): Under OLS 1 – OLS 3,

$$\sqrt{n}(\hat{\beta}_{OLS} - \beta) \xrightarrow{d} N(0, M^{-1} V M^{-1})$$

Thm (Asym. Normality): Under OLS 1 – OLS 5,

$$\sqrt{n}(\hat{\beta}_{OLS} - \beta) \xrightarrow{d} N(0, \sigma^2 M^{-1})$$

3. Instrumental Variables

$$\underset{n\times 1}{y} = \underset{n\times k}{X}\,\underset{k\times 1}{\beta} + \underset{n\times 1}{\epsilon} \quad ; \quad \underset{n\times \ell}{Z} = \begin{pmatrix} z_1' \\ \vdots \\ z_n' \end{pmatrix}, \quad \ell \geq K$$

(IV 1) $\quad Z'X/n \xrightarrow{p} D$, full column rank

$\qquad\qquad Z'Z/n \xrightarrow{p} C$, nonsingular

(IV 2) $\quad Z'\epsilon/n \xrightarrow{p} 0$

(IV 3) $\quad Z'\epsilon/\sqrt{n} \xrightarrow{d} N(0, B)$, $\quad B = Var(Z'\epsilon/\sqrt{n})$

<u>Thm ($L = k$):</u>

Under IV1 - IV3, $\hat{\beta}_{IV} = (Z'X)^{-1}(Z'y) \xrightarrow{p} \beta$

<u>Thm ($L > k$):</u>

Under IV1 - IV3, $\hat{\beta}_{2SLS} = (X'P_Z X)^{-1}(X'P_Z Y) \xrightarrow{p} \beta$

This is 2SLS : Two stage least Squares

Simultaneous Equations Models

Introduction

We have the following Structural Form equations:

$$I: \; y_t = \beta_{10} + \beta_{11} p_t + \epsilon_{1t}$$

$$II: \; p_t = \beta_{20} + \beta_{21} y_t + \beta_{22} I_t + \epsilon_{2t}$$

y_t, p_t are endogenous variables: this gives us the problem of having p_t correlated with ϵ_{1t} and y_t correlated with ϵ_{2t} ∴ OLS on I or II is inconsistent

I_t is our exogenous variable, determined outside the system and independent of $\epsilon_{2t}, \epsilon_{1t}$.

The questions we will be primarly interested in are

1) Which structural form equations are identified, i.e., for which equations can we obtain consistent estimates of the parameters?
2) What restrictions can we impose to obtain identifiability, and what estimation technique do we employ to obtain consistent estimates.

This handout focuses on Question 1.

Eyeballing the structural form equations above, it is obvious that only Eq. (1) is identified. This follows since I_t can be used as an instrument for p_t, while no instrument is available for y_t in Eq. (2).

Solving for y_t, p_t in terms of I_t,

$$y_t = (\beta_{10} + \beta_{11}\beta_{20})/\Delta + \beta_{11}\beta_{22}/\Delta\, I_t + (1/\Delta\, \epsilon_{1t} + \beta_{11}/\Delta\, \epsilon_{2t})$$

$$p_t = (\beta_{20} + \beta_{21}\beta_{10})/\Delta + \beta_{22}/\Delta\, I_t + (\beta_{21}/\Delta\, \epsilon_{1t} + 1/\Delta\, \epsilon_{2t}),$$

$\Delta = 1 - \beta_{11}\beta_{21}$. This can be written in <u>Reduced Form</u>:

$$y_t = \pi_{10} + \pi_{11} I_t + v_{1t}$$

$$p_t = \pi_{20} + \pi_{21} I_t + v_{2t}.$$

The reduced form parameters can be estimated consistently by OLS. Note that $\pi_{11}/\pi_{21} = \beta_{11}$ so that we can estimate β_{11} with $\hat{\pi}_{11,OLS}/\hat{\pi}_{21,OLS}$. Confirming our intuition that Eq. (II) is not identified, no function of the π's gives us back an estimate of β_{21}.

- It should be noted that if we had more than one instrument for p_t (making Eq. (1) <u>over-identified</u>), the reduced form parameters would have yielded an infinite number of estimators for β_{11} (by taking linear combinations of the various available estimators). ✓

- Estimation of the structural form parameters from the reduced form parameters is <u>Indirect Least Squares.</u>

Simultaneas Equations

Let's imagine that observations are made over time so that the i^{th} observation is collected at time i: .

At each i, we observe the row vector

$$\underset{1\times m}{(Y_i},\ \underset{1\times K}{Z_i)}$$

and we imagine that they are related by a system of m equations:

$$\underset{1\times m}{Y_i}\,\underset{m\times m}{B} + \underset{1\times K}{Z_i}\,\underset{K\times m}{\Gamma} = \underset{1\times m}{U_i}, \quad i=1,\dots,n \qquad E(U_i)=0 \quad E(U_i'U_i)=\Sigma$$

This is the <u>Structural Form</u>.

All simultaneous equations models can be written in this way.

<u>Example 1</u>:

$$y_{1t} = \beta_1 y_{2t} + \gamma_1 Z_{2t} + u_{1t}$$
$$y_{2t} = \beta_2 y_{1t} + \gamma_2 Z_{1t} + u_{2t}$$

$$Y_t = (y_{1t}\ \ y_{2t}) \qquad Z_t = (Z_{1t}\ \ Z_{2t}) \qquad U_t = (u_{1t}\ \ u_{2t})$$

$$B = \begin{pmatrix} 1 & -\beta_2 \\ -\beta_1 & 1 \end{pmatrix} \qquad \Gamma = \begin{pmatrix} 0 & -\gamma_2 \\ -\gamma_1 & 0 \end{pmatrix}$$

Y_i and Z_i represent the m endogenous and K exogenous variables, respectively. Z_i, U_i are independent.

Let $A := \begin{bmatrix} B \\ \Gamma \end{bmatrix}$. A collects the structural form parameters. Think of $\begin{bmatrix} B \\ \Gamma \end{bmatrix}$ as m horizontally stacked column vectors.

$$\begin{bmatrix} B \\ \Gamma \end{bmatrix} = \left[\begin{matrix} B_1 \\ \Gamma_1 \end{matrix} \middle| \cdots \middle| \begin{matrix} B_m \\ \Gamma_m \end{matrix} \right]$$

The j^{th} column of $\begin{bmatrix} B \\ \Gamma \end{bmatrix}$ contains the coefficients of the j^{th} equation.

Stacking the observations vertically, we have

$$\begin{matrix} Y_1 B + Z_1 \Gamma = U_1 \\ \vdots \\ Y_n B + Z_n \Gamma = U_n \end{matrix} \quad \Rightarrow \quad \underset{n\times m}{Y} \underset{m\times m}{B} + \underset{n\times K}{Z} \underset{K\times m}{\Gamma} = \underset{n\times m}{U}$$

The i^{th} row of $[Y, Z]$ is the observation $[Y_i, Z_i]$.

<u>Example 2</u>: From example 1, for $t = 1, \dots, T$, we can group Y_t, Z_t so that

$$\underset{T\times 2}{Y} \underset{2\times 2}{B} + \underset{T\times 2}{Z} \underset{2\times 2}{\Gamma} = \underset{T\times 2}{U}$$

. The t^{th} row of $[Y, Z]$ is observation $[Y_t, Z_t] = [y_{1t}\ y_{2t},\ z_{1t}\ z_{2t}]$

The 1st column of $\begin{bmatrix} B \\ \Gamma \end{bmatrix}$ is $\begin{pmatrix} 1 \\ -\beta_1 \\ 0 \\ \gamma_1 \end{pmatrix}$, the coefficients for Eq. (1). ∎

We are now ready to give the <u>Reduced Form</u>

$$Y B + Z \Gamma = U$$

$$Y = -Z \Gamma B^{-1} + U B^{-1}$$

$$\underset{n \times m}{Y} = \underset{n \times k}{Z} \underset{k \times m}{\Pi} + \underset{n \times m}{V}, \qquad \Pi = -\Gamma B^{-1}$$

$$V = U B^{-1}$$

Think back to the reduced form for y_t, p_t given in the introduction. There, y_t and p_t were expressed in terms of the 1 exogenous variable I_t. In this general framework, there are K exogenous variables. The reduced form expresses each of the m endogenous variables in terms of the K exogenous variables.

Look at $\underset{K \times m}{\Pi}$. Π has on each column j the reduced form for the j^{th} equation

Considering only the observation at time i, we look at the i^{th} row of the reduced form:

$$\underset{1 \times m}{Y_i} = \underset{1 \times k}{Z_i} \underset{k \times m}{\Pi} + \underset{1 \times m}{V_i}, \qquad \underset{1 \times m}{V_i} = \underset{1 \times m}{U_i} B^{-1}$$

The covariance matrix for the reduced form errors is $E(V_i'V_i) = E[(U_iB^{-1})'(U_iB^{-1})]$

$$= B^{-1'}E(U_i'U_i)B^{-1}$$

$$= B^{-1'}\Sigma B^{-1} = \Omega$$

<u>Example 3</u>: Continuing examples 1, 2, we have

$$\underset{T\times 2}{Y} = \underset{T\times 2}{Z}\,\underset{2\times 2}{\Pi} + \underset{T\times 2}{V}, \quad \Pi = -\frac{1}{\Delta}\begin{pmatrix} \gamma_2\beta_1 & \gamma_2 \\ \gamma_1 & \gamma_1\beta_2 \end{pmatrix}, \quad \Delta = 1-\beta_1\beta_2$$

Note that if $\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix}$, then indirect least squares estimates of β_1, β_2 would be $\hat{\pi}_{11}/\hat{\pi}_{12}$ and $\hat{\pi}_{22}/\hat{\pi}_{21}$, respectively. Both parameters are <u>exactly identified</u>. ▨

Now we are ready to deal with identification

Recall $A = \begin{bmatrix} B \\ \Gamma \end{bmatrix}$. Consider identification for the first equation, which has parameters $A_1 = \begin{bmatrix} B_1 \\ \Gamma_1 \end{bmatrix}$.

Recall $\underset{K\times m}{\Pi} = -\underset{K\times m}{\Gamma}\,\underset{m\times m}{B^{-1}}$ so that $\underset{K\times m}{\Pi}\,\underset{m\times m}{B} + \underset{K\times m}{\Gamma} = \underset{K\times m}{0}$.

This holds for all m equations, so that $\Pi B_1 + \Gamma_1 = 0$ is also true. Write this as

$$\big[\underset{K\times m}{\Pi} \quad \underset{K\times K}{I_K}\big]\,\underset{(K+m)\times 1}{A_1} = \underset{K\times 1}{0}$$

In addition, we impose g linear restrictions on A_1. We write these as

$$\underset{g\times(m+k)}{\Phi}\ \underset{(m+k)\times 1}{A_1} = \underset{g\times 1}{\phi}\ , \quad \text{where } \Phi \text{ and } \phi \text{ are known.}$$

Then $A_1 = \begin{bmatrix} B_1 \\ \Gamma_1 \end{bmatrix}$ is identified iff

$$\underset{(k+g)\times(m+k)}{\begin{bmatrix} \Pi & I_k \\ \hline \multicolumn{2}{c}{\Phi} \end{bmatrix}} A_1 = \begin{bmatrix} 0 \\ \phi \end{bmatrix}$$

has a unique solution for A_1.

<u>Example 4</u>: Continuing examples 1-3, for equation 1 we have the following restrictions for $A_1 = \begin{bmatrix} B_1 \\ \Gamma_1 \end{bmatrix}$.

Denoting the parameter on y_{1t} by β_0 and the parameter on z_{1t} by γ_0 so that

$$A_1 = \begin{bmatrix} \beta_0 \\ -\beta_1 \\ -\gamma_0 \\ -\gamma_1 \end{bmatrix}, \quad \text{we impose } \beta_0 = 1 \text{ and } \gamma_0 = 0.$$

This gives $\Phi = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, $\phi = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. $\blacksquare$

<u>Thm</u>: A_1 is identified if

$$\text{Rank} \underset{(k+g)\times(m+k)}{\begin{bmatrix} \Pi & I_k \\ \hline \multicolumn{2}{c}{\Phi} \end{bmatrix}} = m+k$$

This theorem will be used to formulate a more precise rank condition. But first note that for $\text{rank}\underset{(k+g)\times(m+k)}{\begin{bmatrix} \Pi & I_k \\ \hline \multicolumn{2}{c}{\Phi} \end{bmatrix}} = m+k$ to have any chance of holding, we need at least $m+k$ ~~columns~~ rows in $\begin{bmatrix} \Pi & I_k \\ \hline \multicolumn{2}{c}{\Phi} \end{bmatrix}$. This gives us the necessary order condition:

Order Condition (1): $g \geq m$

(# linear restrictions) $\geq$ (# endog. variables).

Note that if we normalize so that the β coefficient on the LHS variables in each structural equation are set to 1 (e.g. the coefficient on y_{1t} in Eq. (1) of Example 1 is set to 1; the coefficient on y_{2t} in Eq. (2) of Example 1 is set to 1), then we no longer need to consider this normalization restriction in Φ. Removing this normalization restriction from Φ gives us the order condition: $g^{*}+1 \geq m$

$\uparrow$ normalization

(# linear restrictions without normalization restriction)

In this framework, g^* now includes only <u>exclusion restrictions</u>. These are the restrictions that set certain parameters to 0, thereby excluding the associated variables from the equation.

<u>Example 5</u>

$$
\begin{aligned}
y_{1t} &= \beta_{12} y_{2t} + \gamma_{11} Z_{1t} + \gamma_{12} Z_{2t} + u_{1t} \\
y_{2t} &= \beta_{21} y_{1t} + \beta_{23} y_{3t} + \gamma_{22} Z_{2t} + \gamma_{23} Z_{3t} + u_{2t} \\
y_{3t} &= \beta_{31} y_{1t} + \gamma_{31} Z_{1t} + u_{3t}
\end{aligned}
$$

$$
B_1 = \begin{pmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{13} \end{pmatrix} = \begin{pmatrix} 1 \\ \beta_{12} \\ 0 \end{pmatrix} \qquad \Gamma_1 = \begin{pmatrix} \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \end{pmatrix} = \begin{pmatrix} \gamma_{11} \\ \gamma_{12} \\ 0 \end{pmatrix}
$$

Sign errors

We do not include the normalization restriction in g, the # of exclusion restrictions. For equation 1,

$$
\Phi^* = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \qquad \phi^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
$$

$\Phi^* A_1 = \phi^*$, where Φ^* does not include the normaliz restriction.

$g^* = \#$ exclusion restrictions $= 2$

$m = 3$

$$
\underset{g^*}{\underset{\uparrow}{2}} + \underset{\text{normalization restriction}}{\underset{\uparrow}{1}} \geq 3 \implies \text{Order Condition satisfied for equation 1.}
$$

Now, separating g^* into

g_1^*: # exclusion restrictions for endog. variables

g_2^*: # exclusion restrictions for exog. variables,

we have
$$g_1^* + g_2^* + 1 \geq m$$
$$g_2^* + 1 \geq (m - g_1^*)$$
$$\underbrace{g_2^*}_{\text{\# excluded exogenous variables}} \geq \underbrace{(m - g_1^*)}_{\text{\# of included endogenous variables}} - 1$$

This gives us an alternative way of writing the order condition:

<u>Order Condition (2)</u>: (# excluded exog. variables) $\geq$ (# of included endog. variables) $-$ 1.

Sometimes you'll see this written as

(# excl. exog. var.) $\geq$ (# RHS incl. endog. variables).

Those are the same thing since the right hand side of the two inequalities are the same.

For instance, $y_{1t} = \beta_1 y_{2t} + \gamma_1 z_{1t} + u_t$ gives

included endog. variables = 2 (y_{1t} and y_{2t})

RHS incl. endog. variables = 1 (only y_{2t}).

Example 6: For Example 1, Eq. (1), we have
2 included endog. variables, 1 excluded exog. variable
so $1 \geq 2-1 = 1$, order condition is satisfied.

For Example 5, let r = # included endog. variables
s = # excluded exog. variables

Then Eq. (1): $r=2$, $s=1$ Order Condition Satisfied
Eq. (2): $r=3$, $s=1$ Order Condition Fails
Eq. (3): $r=2$, $s=2$ Order Condition Satisfied

Now lets turn back to the rank condition.

Requiring $\operatorname{rank} \underbrace{\begin{bmatrix} \Pi & I_k \\ \multicolumn{2}{c}{\Phi} \end{bmatrix}}_{(k+g) \times (m+k)} = m+k$ reduces to

requiring $\operatorname{rank}(\Phi A) = m$. Lets see why:

$$\underbrace{\begin{bmatrix} \overset{k\times m}{\Pi} & \overset{k \times k}{I} \\ \multicolumn{2}{c}{\underset{g\times(m+k)}{\underbrace{\Phi}}} \end{bmatrix}}_{(k+g)\times(m+k)} \underbrace{\begin{bmatrix} \overset{m\times m}{B} & 0 \\ \underset{k\times m}{\Gamma} & \underset{k\times k}{I} \end{bmatrix}}_{\text{nonsingular}} = \begin{bmatrix} 0 & I \\ \Phi A & \Phi \begin{bmatrix} 0 \\ I \end{bmatrix} \end{bmatrix}$$

Since the second matrix on the LHS is nonsingular, the product of the two matrices has the same rank as the first matrix on the LHS.

Now, $\operatorname{rank}\begin{pmatrix} \Pi & I \\ \Phi & \end{pmatrix} = \operatorname{rank}\begin{pmatrix} 0 & I \\ \Phi A & \Phi\left[\begin{smallmatrix} 0 \\ I \end{smallmatrix}\right] \end{pmatrix} = \operatorname{rank}(\Phi A) + k$

<u>Rank Condition</u>: $\operatorname{rank}(\Phi A) = m$

Identifiability:

- If the order or rank condition fails, the equation is <u>not identified</u>.
- If the order condition holds with equality, the equation is <u>just identified</u> if the rank condition holds.
- If the order condition holds with inequality, and the rank condition holds, the equation is <u>over-identified</u>.

The order condition is a necessary condition for identifiability.

The rank condition is necessary and sufficient.

Example 7: Discuss identifiability of Example 5.

Here, $A = \begin{pmatrix} 1 & \beta_{21} & \beta_{31} \\ \beta_{12} & 1 & 0 \\ 0 & \beta_{23} & 1 \\ \gamma_{11} & 0 & \gamma_{31} \\ \gamma_{12} & \gamma_{22} & 0 \\ 0 & \gamma_{23} & 0 \end{pmatrix}$

Eq. (1): From Ex. 6, order condition satisfied with equality.

normalization

$$\Phi = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}; \quad \Phi A = \begin{pmatrix} 1 & \beta_{21} & \beta_{31} \\ 0 & \beta_{23} & 1 \\ 0 & \gamma_{23} & 0 \end{pmatrix}$$

restriction restriction

$\text{rank}(\Phi A) = 3 = m$ for $\gamma_{23} \neq 0$.

Eq. (1) is just-identified.

Eq. (2): From Ex. 6, order condition fails. No need to check

Eq. (2) is not-identified.

Eq. (3): $\Phi = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}; \quad \Phi A = \begin{pmatrix} \beta_{12} & 1 & 0 \\ 0 & \beta_{23} & 1 \\ \gamma_{12} & \gamma_{22} & 0 \\ 0 & \gamma_{23} & 0 \end{pmatrix}$

$\text{rank}(\Phi A) = 3 = m$

From Ex. 6, order condition holds with inequality.

Eq. (3) is over-identified. $\blacksquare$.