EC402 (2023/24)

RAGVIR'S SPEAKING NOTES FOR CLASS

Vassilis PS 8 (WK 9, AT)

QUESTION 1

2018

The NLRM is as follows:

A1
A2
A3R mi (or stronger)
A4GM
A5N

$$y = X\beta + \varepsilon$$

- $\varepsilon \sim N(0, \sigma^2 I_N)$
- $E[\varepsilon | X] = E[\varepsilon]$
- X is st $(X'X)^{-1}$ exists

Then, $\hat{\beta}_{OLS} | X \sim MVN\left(\beta, \sigma^2 (X'X)^{-1}\right)$

Alternatively,

A1
A2
A3R mi (or stronger)
A4 Ω
A5N

$$y = X\beta + \varepsilon$$

- $\varepsilon \sim N(0, \Omega)$
- $E[\varepsilon | X] = E[\varepsilon]$
- X is s.t. $(X'X)^{-1}$ exists

Then, $\hat{\beta}_{OLS} | X \sim MVN\left(\beta, (X'X)^{-1} X' \Omega X (X'X)^{-1}\right)$

2018

The Asymptotic NLRM (ANLRM):

A1
A2
A3 R srs (or stronger)
A4 GM

$$y = X\beta + \varepsilon$$

- $\varepsilon \sim$?? $(0, \sigma^2 I_N)$
- $E[\varepsilon_i x_{hi}] = 0$ for all $i = 1, \dots, N;\ h = 1, \dots, K$
- $(X'X)^{-1}$ exists and $\left(\frac{X'X}{N}\right) \xrightarrow{p} Q$ a P.D matrix

Then $\sqrt{N}\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} MVN\left(0, \sigma^2 Q^{-1}\right)$

as $N \to \infty$.

A1
A2
A3 R srs (or stronger)
A4 Ω

$$y = X\beta + \varepsilon$$

- $\varepsilon \sim$?? $(0, \Omega)$
- $E[\varepsilon_i x_{hi}] = 0$ for all $i = 1, \dots, N;\ h = 1, \dots, K$
- $(X'X)^{-1}$ exists and $\left(\frac{X'X}{N}\right) \xrightarrow{p} Q$ a P.D. matrix

Then, $\sqrt{N}\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} MVN\left(0, Q^{-1} Q^{*} Q^{-1}\right)$ as $N \to \infty$

where $Q^{*} = \text{plim}\left(X'\Omega X / N\right)$ a P.D matrix.

2018

NLRM (w. A4GM)

(i) $R\hat{\beta}_{OLS} | X \sim MVN(R\beta, R\,\sigma^2 (X'X)^{-1} R')$

(ii) $\hat{\sigma}^2_{OLS} := (\hat{\varepsilon}'_{OLS}\hat{\varepsilon}_{OLS})/(N-K)$; $E[\hat{\sigma}^2_{OLS}|X] = \sigma^2$ & $\frac{(N-K)\hat{\sigma}^2_{OLS}}{\sigma^2} \Big| X \sim \chi^2_{N-K}$

(iii) $(R\hat{\beta}_{OLS} - R\beta)'(R'\,\sigma^2 (X'X)^{-1} R')^{-1}(R\hat{\beta}_{OLS} - R\beta) \,|\, X \sim \chi^2_r$

(iva) Conditional on X, the RVs in (i) and (ii) are independent.

(ivb) Conditional on X, the RVs in (ii) and (iii) are independent.

(v) Conditional on X, $T_j = (\hat{\beta}^j_{OLS} - \beta^j) \big/ \sqrt{[\hat{\sigma}^2_{OLS}(X'X)^{-1}]_{jj}} \sim t_{N-K}$

(vi) Conditional on X, $F = (R\hat{\beta}_{OLS} - R\beta)'[R\,\hat{\sigma}^2_{OLS}(X'X)^{-1} R']^{-1}(R\hat{\beta}_{OLS} - R\beta)/r \sim F_{r,N-K}$

Finally, ask yourself how above results would be modified under the ANLRM (w. A4GM).

[P.S. Also see Vassilis' ~~solutions~~ for OLS Results (or lack thereof) in the NLRM|ANLRM (w. A4Ω)]

TIP: TRY TO STUDY PAGES 94-96 OF VASSILIS' "SUPPLEMENTARY TECHNICAL NOTES" FOR AN UNDERSTANDING OF WHAT YOU'RE SUPPOSED TO DISCUSS.

I WILL GUIDE YOU (THROUGH PAGE 94) IN CLASS (SEE SLIDE BELOW).

CASE I:	**(A)NLRM with A4GM(iid)**
Result 1	Conditionally on X, $R\hat{\beta}_{OLS}\|X \sim N\left(R\beta^{true}, R\left[\sigma^2_{\epsilon^{true}}(X'X)^{-1}\right]R'\right)$
Result 2	$\hat{s}^2_{OLS} \equiv \frac{RSS_{OLS}}{S-k}$ is unbiased for $\sigma^2_{\epsilon^{true}}$ & conditionally on X, $\frac{(S-K)\hat{s}^2_{OLS}}{\sigma^2_{\epsilon^{true}}} \sim \chi^2(S-k)$
Result 3	Conditionally on X, $(R\hat{\beta}_{OLS} - R\beta^{true})'\left(R\left[\sigma^2_{\epsilon^{true}}(X'X)^{-1}\right]R'\right)^{-1}(R\hat{\beta}_{OLS} - R\beta^{true}) \sim \chi^2(r)$
Result 4A	Conditionally on X, the Gaussian r.v. $R\hat{\beta}_{OLS}$ of Result 1 and the χ^2 r.v. $\frac{(S-K)\hat{s}^2_{OLS}}{\sigma^2_{\epsilon^{true}}}$ of Result 2 are *independent*
Result 4B	Conditionally on X, theχ^2 r.v. $\frac{(S-K)\hat{s}^2_{OLS}}{\sigma^2_{\epsilon^{true}}}$of Result 2 and the χ^2 r.v. of Result 3 are *independent*
Result 5	Conditionally on X, $\tau \equiv \frac{\hat{\beta}^j_{OLS}-\beta^j_{true}}{\sqrt{\left[\hat{s}^2_{OLS}(X'X)^{-1}\right]_{jj}}} \sim t(S-k)$ because of Results 1,2,&4A
Result 6	Conditionally on X, $f \equiv (R\hat{\beta}_{OLS} - R\beta^{true})'\left(R\left[\hat{s}^2_{OLS}(X'X)^{-1}\right]R'\right)^{-1}(R\hat{\beta}_{OLS} - R\beta^{true})/r \sim F(r, S-k)$
Result 7	Suppose an estimator $\hat{\theta}_{method}$ for the $p \times 1$ unknown parameter vector θ^{true} obeys: $\hat{\theta}_{method} \sim N(\theta^{true}, V_{GM}(\hat{\theta}_{method}))$ for any sample size S; or $\hat{\theta}_{method} \approx N(\theta^{true}, V_{GM}(\hat{\theta}_{method}))$ for very large sample size S. Then for any continuous function $g(.) : R^p \to R^r$ with continuous first-derivative matrix $\left[\frac{\partial g(.)}{\partial \theta}\right]$ it follows that: $g(\hat{\theta}_{method}) \approx N\left(g(\theta^{true}), \left[\frac{\partial g(.)}{\partial \theta}\right] V_{GM}(\hat{\theta}_{method}) \left[\frac{\partial g(.)}{\partial \theta}\right]'\right)$ for very large S.

Question 2

2018

(Q2) let $y = X\beta + \varepsilon$ where $\underset{N\times K}{X} = [x_1 \; x_2 \ldots x_K]$ and $\hat{\beta} := (X'X)^{-1}X'y$

Let's assume: A1, A2, A3R srn, A4GM.

NON-LINEAR RESTRICTIONS RECAP:

- let $g(\beta) \in \mathbb{R}^r$ be a set of r non-linear and continuously differentiable functions of $\beta \in \mathbb{R}^K$.

Say we wish to test $H_0: g(\beta) = 0$

$H_1: g(\beta) \neq 0$

- Using the WALD PRINCIPLE of testing based on $\hat{\beta} := (X'X)^{-1}X'y$ (i.e. from the unrestricted model), we would ideally wish to construct a statistic of the form:

$$[g(\hat{\beta})]' \left[\hat{Var}(g(\hat{\beta}))\right]^{-1} [g(\hat{\beta})]$$

as the non-linear analogue to the usual $(R\hat{\beta}-q)' [R \hat{Var}(\hat{\beta})R']^{-1} (R\hat{\beta}-q) \overset{approx.}{\sim} \chi^2_r$ statistic as in the case of r linear restrictions contained in the matrix R.

The problem is that we need $\hat{var}(g(\hat{\beta}))$ where $g(.)$ is a NON-LINEAR function.

(i.e. you can't just "pull out" the "R" happily "with a square"!)

2018

- We use the result (from the "delta method") that:

For "large" N (under A1, A2, A3R sru, A4GM),

$$g(\hat\beta) \overset{approx}{\sim} N\Big(g(\beta), \underbrace{[\partial g(\beta)/\partial\beta']}_{\text{Kinda like an "R"}} Var(\hat\beta) \underbrace{[\partial g(\beta)/\partial\beta']'}_{\text{..."R'"}}\Big).$$

- We can proceed then with test statistic V:

$$V := g(\hat\beta)' \Big[[\partial g(\hat\beta)/\partial\hat\beta'] [\hat\sigma^2 (X'X)^{-1}] [\partial g(\hat\beta)/\partial\hat\beta']' \Big]^{-1} g(\hat\beta)$$

Under H_0: $g(\beta)=0$, $V \overset{approx.}{\sim} \chi^2_r$ for "large" N (assuming A1, A2, A3R sru, A4GM).

- For Q2, we have in particular:

$$g(\beta) := \begin{bmatrix} \beta_1\beta_2 - 1 \\ \beta_3 - 4\beta_4 + 2 \end{bmatrix} \text{ and } \partial g(\beta)/\partial\beta' := \begin{bmatrix} \beta_2 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 \end{bmatrix}$$

Thus, $V = [\hat\beta_1\hat\beta_2 - 1, \hat\beta_3 - 4\hat\beta_4 + 2] \left[\begin{pmatrix} \hat\beta_2 & \hat\beta_1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 \end{pmatrix} \hat\sigma^2 (X'X)^{-1} \begin{pmatrix} \hat\beta_2 & 0 \\ \hat\beta_1 & 0 \\ 0 & 1 \\ 0 & -4 \\ 0 & 0 \end{pmatrix} \right]^{-1} \begin{bmatrix} \hat\beta_1\hat\beta_2 - 1 \\ \hat\beta_3 - 4\hat\beta_4 + 2 \end{bmatrix}$

$\overset{approx.}{\sim} \chi^2_2$ under H_0 for "large N" (assuming A1, A2, A3R sru, A4GM).

- We can thus find for some exogenously set significance level $\alpha \in (0,1)$ a critical value, say $\chi^2_{2,1-\alpha}$, which denotes the $100(1-\alpha)^{th}$ percentile on the CDF of a χ^2_2 random variable.

- Our decision rule would be to:

  Reject H_0 at the $100\alpha\%$ significance level if and only if $V > \chi^2_{2,1-\alpha}$.

- Finally, we note that a distributional (A5 style) assumption is not needed since the Delta method only gives us asymptotically valid distributional results for V. No exact results are available for V in finite samples.

[Test yourself: can you spot the typos in the official solution? Hint: K=5.]

- An alternative method may be:

① estimate the unrestricted model via MLE (assuming ASN) ← why?!

i.e. $\tilde{\beta} := \arg \underset{\beta}{\cancel{\min}^{\max}} \log \mathcal{L}(\beta; \text{data})$

and compute $LLF_u := \log \mathcal{L}(\tilde{\beta}; \text{data})$

② estimate the restricted model via MLE (assuming ASN) why?!

i.e. $\tilde{\beta}_R := \arg \underset{\beta}{\cancel{\min}^{\max}} \log \mathcal{L}(\beta; \text{data})$

s.t. $g(\beta) = 0$

and compute $LLF_R := \log \mathcal{L}(\tilde{\beta}_R; \text{data})$

③ $LR := 2(LLF_u - LLF_R) \overset{\text{approx.}}{\sim} \chi^2_r$ for "large" N.

Now, do the usual "recipe" here. (eg: Set α; Find $LR_{crit, 1-\alpha}$; Define rejection region, etc)

... to finish your solution.

Found this; thought it may help you. It's what I would have also said to explain the previous slide.

2020

23:06 Wed 25 Nov

81%

Back

Marno Verbeek-A Guide to Modern Econometrics-Wiley (2012).pdf

The **likelihood ratio test** is even simpler to compute, provided the model is estimated with and without the restrictions imposed. This means that we have two different estimators: the unrestricted ML estimator $\hat{\theta}$ and the constrained ML estimator $\tilde{\theta}$, obtained by maximizing the loglikelihood function $\log L(\theta)$ subject to the restrictions $R\theta = q$. Clearly, maximizing a function subject to a restriction will not lead to a larger maximum compared with the case without the restriction. Thus it follows that $\log L(\hat{\theta}) - \log L(\tilde{\theta}) \geq 0$. If this difference is small, the consequences of imposing the restrictions $R\theta = q$ are limited, suggesting that the restrictions are correct. If the difference is large, the restrictions are likely to be incorrect. The LR test statistic is simply computed as

$$\xi_{LR} = 2[\log L(\hat{\theta}) - \log L(\tilde{\theta})],$$

which, under the null hypothesis, has an approximate Chi-squared distribution with J degrees of freedom. This shows that, if we have estimated two specifications of a model, we can easily test the restrictive specification against the more general one by comparing loglikelihood values. It is important to stress that the use of this test is only appropriate if the two models are nested (see Chapter 3). ~~An attractive feature of the test is that it is~~

QUESTION 3

$y = X\beta + \varepsilon$; let X be a $T \times k$ matrix

A1, A2, A3Rmi hold. Also, A4 Ω. let $\Omega \equiv \Omega(\lambda)$,
& $\hat{\Omega} \equiv \Omega(\hat{\lambda})$.

(a) $\hat{\beta}_{GLS} := (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$

$\hat{\beta}_{FGLS} := (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y$

Common exam errors: forgetting "-1" symbols; writing $(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}X_y$ forgetting hats; talking about θ, when parameter is λ.

(b) $\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y \overset{A2}{=} (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}(X\beta + \varepsilon)$

$= (X'\Omega^{-1}X)X'\Omega^{-1}X\beta + (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\varepsilon$

$= \beta + (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\varepsilon$

Similarly, $\hat{\beta}_{FGLS} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y \overset{A2}{=} (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}(X\beta + \varepsilon)$

$= \beta + (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}\varepsilon$

∴ The sampling error vectors in each case are as follows:

for IGLS case : $(\hat{\beta}_{GLS} - \beta) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\varepsilon$

for FGLS case : $(\hat{\beta}_{FGLS} - \beta) = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}\varepsilon$

(c) Properties of $(\hat{\beta}_{GLS} - \beta)$: [1] Unbiasedness :

- $E[\hat{\beta}_{GLS} - \beta] = E[(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\varepsilon] \overset{UE}{=} E\left[E[(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\varepsilon \mid X, \Omega]\right]$

treat as known

$\overset{\text{Linearity}}{=} E\left[(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} E[\varepsilon \mid X, \Omega]\right]$

$= E\left[(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} E[\varepsilon \mid X]\right] \overset{A3Rmi}{=} 0$

∴ $E[\hat{\beta}_{GLS}] - \beta = 0$ so $\hat{\beta}_{GLS}$ is an unbiased estimator of β.

- $E\left[\hat{\beta}_{FGLS} - \beta\right] = E\left[(X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}\varepsilon\right] \overset{\text{LIE}}{=} E\left[E\left[(X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}\varepsilon \,\middle|\, X, \Omega\right]\right]$

$$= E\left[E\left[(X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}\varepsilon \,\middle|\, X\right]\right]$$

Recall: $\hat{\Omega} = \Omega(\hat{\lambda})$ where $\hat{\lambda}$ is an estimator of λ, some parameter which controls/define the structure of Ω. Thus, $\hat{\lambda} := g(y, X) = g_2(\varepsilon, \beta, X)$ makes $\hat{\Omega}$ and, in turn, $(\hat{\beta}_{FGLS} - \beta)$ potentially a non-linear function of ε. So, we cannot rely, as we usually might, on linearity property of the $E[\cdot|X]$ operator to continue the unbiasedness proof.

It is, in general, not known how to find unbiasedness of $\hat{\beta}_{FGLS}$ as an estimator of β.

(Indeed, finite-sample properties and exact distributions of FGLS estimators are generally not known.)

$\therefore$ COMPARISON I: $\hat{\beta}_{GLS}$ is an unbiased estimator of β under A1, A2, A3Rnd and A4Ω but the same cannot be said in general for $\hat{\beta}_{FGLS}$.

[2] Consistency (as $T \to \infty$):

- $$\operatorname*{plim}_{T\to\infty} (\hat{\beta}_{GLS} - \beta) = \operatorname*{plim}_{T\to\infty}\left[\left(\frac{1}{T}X'\Omega^{-1}X\right)^{-1}\frac{1}{T}X'\Omega^{-1}\varepsilon\right]$$
$$= \left(\operatorname*{plim}_{T\to\infty}\frac{1}{T}(X'\Omega^{-1}X)\right)^{-1}\cdot \operatorname*{plim}_{T\to\infty}\frac{1}{T}(X'\Omega^{-1}\varepsilon)$$
by Slutsky's theorem.

- Now, assume G1: $\operatorname*{plim}_{T\to\infty} \frac{1}{T}(X'\Omega^{-1}X) = \underset{K\times K}{Q^*}$ some finite P.D. matrix

& G2: $\operatorname*{plim}_{T\to\infty} \frac{1}{T}(X'\Omega^{-1}\varepsilon) = 0_{K\times 1}$

$\therefore$ Under A1-A4Ω and G1, G2, $\hat{\beta}_{GLS}$ is a consistent estimator of β (as $T\to\infty$).

$$\bullet\quad \hat\beta_{FGLS}-\beta = (\hat\beta_{FGLS}-\hat\beta_{GLS}) + (\hat\beta_{GLS}-\beta)$$

$$= \left[(X'\hat\Omega^{-1}X)^{-1}X'\hat\Omega^{-1} - (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\right]y + (\hat\beta_{GLS}-\beta)$$

$$= \left[(X'\hat\Omega^{-1}X)^{-1}X'\hat\Omega^{-1} - (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\right]\varepsilon + (\hat\beta_{GLS}-\beta)$$

$$= (X'\hat\Omega^{-1}X)^{-1}X'\hat\Omega^{-1}\varepsilon - (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\varepsilon + (\hat\beta_{GLS}-\beta)$$

$$\therefore \operatorname*{plim}_{T\to\infty}(\hat\beta_{FGLS}-\beta) = \left[\operatorname*{plim}_{T\to\infty}\left(\frac{1}{T}(X'\hat\Omega^{-1}X)\right)\right]^{-1}\operatorname*{plim}_{T\to\infty}\left[\frac{1}{T}X'\hat\Omega^{-1}\varepsilon\right]$$

$$- \left[\operatorname*{plim}_{T\to\infty}\left(\frac{1}{T}(X'\Omega^{-1}X)\right)\right]^{-1}\operatorname*{plim}_{T\to\infty}\left[\frac{1}{T}X'\Omega^{-1}\varepsilon\right]$$

$$+ \operatorname*{plim}_{T\to\infty}(\hat\beta_{GLS}-\beta)$$ by Slutsky's theorem.

COMPARISON 2:

if we have that F1: $\underset{T\to\infty}{\text{plim}}\left[\frac{1}{T}(X'\hat{\Omega}^{-1}X)\right] = \underset{T\to\infty}{\text{plim}}\left[\frac{1}{T}(X'\Omega^{-1}X)\right]$

& F2: $\underset{T\to\infty}{\text{plim}}\left[\frac{1}{T}(X'\hat{\Omega}^{-1}\varepsilon)\right] = \underset{T\to\infty}{\text{plim}}\left[\frac{1}{T}(X'\Omega^{-1}\varepsilon)\right]$ then

under A1-A4Ω and G1, G2 and F1, F2, both $\hat{\beta}_{GLS}$ and $\hat{\beta}_{FGLS}$ are consistent (as $T\to\infty$) estimators of β.

[In practice, we don't "ASSUME" F1 and F2; we prove them based on $\sqrt{T}$-consistency $\hat{\Omega}$ for Ω. But that's not my goal for this problem set. (i.e. in research, it's bad practice to just assume that estimators are consistent)]

[3] Limiting distributions (as $T \to \infty$):

Result:

- Assume G3: $\frac{1}{\sqrt{T}} X'\Omega^{-1}\varepsilon \xrightarrow{d} MVN\left(0, Q^*\right)$ as $T \to \infty$.

$\Rightarrow$ we have that $\sqrt{T}\left(\hat{\beta}_{GLS} - \beta\right) \xrightarrow{d} MVN\left(0, \operatorname*{plim}_{T\to\infty} c^2\left(\frac{X'\Omega^{-1}X}{T}\right)^{-1}\right)$ as $T \to \infty$

- Also assume F3: $\operatorname*{plim}_{T\to\infty} \frac{1}{\sqrt{T}}\left(X'\hat{\Omega}^{-1}\varepsilon\right) = \operatorname*{plim}_{T\to\infty} \frac{1}{\sqrt{T}}\left(X'\Omega^{-1}\varepsilon\right)$

$\Rightarrow$ we have that $\sqrt{T}\left(\hat{\beta}_{FGLS} - \beta\right) \xrightarrow{d} MVN\left(0, \operatorname*{plim}_{T\to\infty} c^2\left(\frac{X'\Omega^{-1}X}{T}\right)^{-1}\right)$ as $T \to \infty$.

COMPARISON 3: Under A1-A4 Ω and G1, G2, G3 and F1, F2, F3, $\hat{\beta}_{FGLS}$ and $\hat{\beta}_{GLS}$ are both asymptotically Normal as $T \to \infty$.