EC402 (2023/24)

RAGVIR'S SPEAKING NOTES FOR CLASS

Vassilis PS 9 (WK10, AT)

QUESTIONS 0 and 1

(RAGVIR'S ADDITION) (VASSILIS' ORIGINAL QUESTION)

Consider $y = X\beta + \varepsilon$, where the t^{th} element of ε is such that $\varepsilon_t \overset{IID}{\sim} (0, \sigma_\varepsilon^2)$ for $t = 1, \dots, T$

and finite σ_ε^2. Assume further that $T \times K$ matrix X has full rank $K \ll T$,

that the rows $\{x_t'\}_{t=1}^{T}$ constitute an IID sequence, and $X \perp\!\!\!\perp \varepsilon$.

(Q0a) Find the probability limit of $(\hat{\beta}_{OLS} - \beta)$;

(Q0b) Find the limiting distribution of $\sqrt{T}(\hat{\beta}_{OLS} - \beta)$; and

(Q1) Find the probability limit of $(\hat{\sigma}_\varepsilon^2 - \sigma^2)$;

as $T \to \infty$, where $\hat{\beta}_{OLS} := (X'X)^{-1}X'y$ and $\hat{\sigma}_\varepsilon^2 := \dfrac{\hat{\varepsilon}'\hat{\varepsilon}}{T-K}$ for $\hat{\varepsilon} := y - X\hat{\beta}_{OLS}$

(Q0 a) PROPOSITION. $(\hat{\beta}_{OLS} - \beta) \xrightarrow{p} 0$ as $T \to \infty$.

PROOF. We have $\hat{\beta}_{OLS} - \beta = (X'X)^{-1} X'\varepsilon = \left(\frac{1}{T}\sum_{t=1}^{T} x_t x_t'\right)^{-1} \frac{1}{T}\sum_{t=1}^{T} x_t \varepsilon_t$

So that $\operatorname{plim}_{T\to\infty} (\hat{\beta}_{OLS} - \beta) = \left(\operatorname{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} x_t x_t'\right)^{-1} \operatorname{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} x_t \varepsilon_t$ by Slutsky's theorem.

Let us add the assumption that $E[x_t x_t']$ is a finite P.D. matrix.

Then, by Khinchine's LLN, it holds that $\operatorname{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} x_t x_t' = E[x_t x_t']$

and by Chebyshev's LLN, it holds that $\operatorname{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} x_t \varepsilon_t = E[x_t \varepsilon_t]$

So that $\operatorname{plim}_{T\to\infty} (\hat{\beta}_{OLS} - \beta) = \left[E[x_t x_t']\right]^{-1} E[x_t \varepsilon_t]$.

We recognise that $E(x_t \varepsilon_t) \overset{LIE}{=} E(x_t E[\varepsilon_t | X]) = 0$ under our modelling framework

and this completes the proof. $\square$

(Q0b) **Proposition.** $\sqrt{T}\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} N\left(0, \sigma_\varepsilon^2 \, E\left[x_t x_t'\right]^{-1}\right)$ as $T \to \infty$.

Proof. We have $\sqrt{T}\left(\hat{\beta}_{OLS} - \beta\right) = \left(\frac{1}{T}\sum_{t=1}^{T} x_t x_t'\right)^{-1} \sqrt{T}\,\frac{1}{T}\sum_{t=1}^{T} x_t \varepsilon_t$.

As noted earlier, it holds by Khinchine's LLN that

$$\operatorname*{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} x_t x_t' = E\left[x_t x_t'\right].$$

Moreover, it holds by the Lindeberg-Lévy CLT that

$$\sqrt{T}\,\frac{1}{T}\sum_{t=1}^{T} x_t \varepsilon_t \xrightarrow{d} N\left(0, \sigma_\varepsilon^2 \, E\left[x_t x_t'\right]\right) \text{ as } T \to \infty.$$

The required result then follows from Cramér's convergence theorem.

$\square$

(Q1) Step 1 : Algebra

Recognise $y = X\beta + \varepsilon$

or $y = X\hat{\beta} + \hat{\varepsilon}$,

So $X(\hat{\beta} - \beta) + (\hat{\varepsilon} - \varepsilon) = 0$

$\Rightarrow \quad \hat{\varepsilon} = \varepsilon - X(X'X)^{-1}X'\varepsilon.$

Thus,

$$\hat{\varepsilon}'\hat{\varepsilon} = \left(\varepsilon - X(X'X)^{-1}X'\varepsilon\right)'\left(\varepsilon - X(X'X)^{-1}X'\varepsilon\right)$$

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

$$= \varepsilon'\varepsilon - \varepsilon'X(X'X)^{-1}X'\varepsilon.$$

Alternatively, we obtain

$$\hat{\varepsilon}'\hat{\varepsilon} = (M_X\varepsilon)'(M_X\varepsilon) = \varepsilon' M_X \varepsilon$$

$$= \varepsilon'\varepsilon - \varepsilon' X(X'X)^{-1}X'\varepsilon.$$

Finally,

$$\frac{\hat{\varepsilon}'\hat{\varepsilon}}{T} = \left[\frac{\varepsilon'\varepsilon}{T} - \left(\frac{X'\varepsilon}{T}\right)'\left(\frac{X'X}{T}\right)^{-1}\left(\frac{X'\varepsilon}{T}\right)\right].$$

Step 2: Assumptions

Assume $E(x_t \varepsilon_t) = 0$, and [Assume X is zero-mean WLOG]

Σ_{XX} is a finite P.D. matrix, and

$\sigma_\varepsilon^2 \in (0, \infty)$,

noting that my exogeneity assumption is of the weakest possible form

Step 3: Slutsky's theorem

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

by repeated use of Slutsky's theorem.

Step 4 : "Suitable" LLN's

(i) By Khinchine's WLLN (ε_t^2 IID with finite mean, as assumed in Step 2, σ_ε^2),

$$\operatorname*{plim}_{T\to\infty}\left(\frac{\varepsilon'\varepsilon}{T}\right) = \operatorname*{plim}_{T\to\infty}\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2\right] = E\left[\varepsilon_t^2\right] = \sigma_\varepsilon^2 .$$

(ii) By Khinchine's WLLN ($x_t x_t'$ IID with finite 1st mom., as assumed in Step 2, Σ_{XX}),

$$\operatorname*{plim}_{T\to\infty}\left(\frac{X'X}{T}\right) = \operatorname*{plim}_{T\to\infty}\left[\frac{1}{T}\sum_{t=1}^{T}x_t x_t'\right] = E\left[x_t x_t'\right] = \Sigma_{XX}.$$

(iii) By Chebyshev's WLLN, ($x_t\varepsilon_t$ IID with finite mean, as assumed in Step 2, 0; AND finite variance, $\sigma_\varepsilon^2\Sigma_{XX}$),

$$\operatorname*{plim}_{T\to\infty}\left(\frac{X'\varepsilon}{T}\right) = \operatorname*{plim}_{T\to\infty}\left[\frac{1}{T}\sum_{t=1}^{T}x_t\varepsilon_t\right] = E\left[x_t\varepsilon_t\right] = 0$$

[In fact, we have that $\frac{X'\varepsilon}{T} \xrightarrow{M.S.} 0$, a much stronger result.]

Step 5 : Consistency proof

By Step 3 and Step 4,

$$\underset{T\to\infty}{\text{plim}} \left[\frac{\hat{\varepsilon}'\hat{\varepsilon}}{T}\right] = \sigma_\varepsilon^2 .$$

Step 6 : Final answer

Noting that $\lim\limits_{T\to\infty} \left[T/(T-k)\right] = 1$, and with a final (unnecessary / trivial) application of Slutsky's theorem,

$$\underset{T\to\infty}{\text{plim}} \left(s^2 \right] = \underset{T\to\infty}{\text{plim}} \left(\frac{\hat{\varepsilon}'\hat{\varepsilon}}{T-k} \right] \overset{(p)}{=} \lim_{T\to\infty} \left(\frac{T}{T-k} \right| \underset{T\to\infty}{\text{plim}} \left(\frac{\hat{\varepsilon}'\hat{\varepsilon}}{T} \right| = \sigma_\varepsilon^2 .$$

QUESTION 2

Intuition behind $\hat{\beta}_{NLLS}$:

Consider for $s = 1, \dots, S$, the model for latent variable $y_s^* = x_s'\beta + \varepsilon_s$, $\varepsilon_s \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$, $\varepsilon \perp X$.

It is impossible to estimate this model since y_s^* is unobserved. We observe instead

$$y_s = \begin{cases} y_s^*, & y_s^* > 0 \\ 0, & y_s^* \le 0 \end{cases}$$

but $E_{(\beta', \sigma^2)}[y_s | X]$ is not linear. We have a standard regression model for y_s, except one in which all neg. values are mapped to 0 (or -999 in Vassilis' solution).

Indeed

$$E_{(\beta',\sigma^2)}[y_s \mid x] = E_{(\beta',\sigma^2)}[y_s \mid y_s > 0, x]\, P_{(\beta',\sigma^2)}(y_s > 0)$$

$$+ \; E_{(\beta',\sigma^2)}[y_s \mid y_s = 0, x]\, P_{(\beta',\sigma^2)}(y_s = 0)$$

$$= \left[x_s'\beta + \sigma \frac{\phi(x_s'\beta/\sigma)}{\Phi(x_s'\beta/\sigma)} \right] \cdot \Phi(x_s'\beta/\sigma)$$

$$+ \; 0 \cdot \left[1 - \Phi(x_s'\beta/\sigma) \right] \quad \text{by Result 1.}$$

(see orange section below if curious.)

REMARK 1.

We note in the previous expression that

$$P_{(\beta',\sigma^2)}(y_s > 0 \mid x_s) = P_{(\beta',\sigma^2)}(y_s^* > 0 \mid x_s)$$

$$= P_{(\beta',\sigma^2)}(x_s'\beta + \varepsilon_s > 0 \mid x_s) = P_{(\beta',\sigma^2)}(\varepsilon_s > -x_s'\beta \mid x_s)$$

$$= P_{(\beta',\sigma^2)}(\varepsilon_s/\sigma > -x_s'\beta/\sigma \mid x_s)$$

$= \Phi(x_s'\beta/\sigma)$ where $\Phi(a)$, $a \in \mathbb{R}$, denotes $P(Z \leq a)$ for $Z \sim N(0,1)$, and using the property that the $N(0,1)$ PDF is an even function. $\square$

To summarise, let

$$g(x_s;\beta) := \underset{(\beta',\sigma^2)}{E}[y_s \mid X] = \left[x_s'\beta + \sigma\frac{\phi(x_s'\beta/\sigma)}{\Phi(x_s'\beta/\sigma)}\right]\Phi(x_s'\beta/\sigma)$$

so that

$$(\hat{\beta}', \hat{\sigma}^2)_{NLS} := \underset{(\beta',\sigma^2)}{\operatorname{argmin}} \sum_{s=1}^{S} \left[\left(y_s - g(x_s;\beta)\right)^2\right],$$ or indeed

$\hat{\beta}_{NLS}$ is s.t.

$$\sum_{s=1}^{S} g^{\beta}(x_s;\hat{\beta}_{NLS})\left[y_s - g(x_s;\hat{\beta}_{NLS})\right] = 0. \quad \text{— } (1)$$

Properties of $\hat{\beta}_{NLLS}$:

The system of equations in (1) is non-linear. No closed-form expression for $\hat{\beta}_{NLLS}$ exists. We use a 1st order Taylor approximation as follows:

$$0 \approx \sum_{s=1}^{S} g^{\beta}(x_s;\beta)(y_s - g(x_s;\beta)) + \sum_{s=1}^{S}\left[g^{\beta\beta'}(x_s;\beta)(y_s - g(x_s;\beta)) - g^{\beta}(x_s;\beta)\,g^{\beta'}(x_s;\beta)\right](\hat{\beta}_{NLLS} - \beta)$$

Thus, under twice-continuous differentiability of $g(x_s; \beta)$ and mean independence of X with ε, we can use Slutsky's theorem together with an LLN in order to establish (as in Q0) consistency of $\hat{\beta}_{NLLS}$. Further, we can use Cramér's theorem together with a CLT in order to establish asymptotic normality for $\sqrt{S}(\hat{\beta}_{NLLS} - \beta)$ as $S \to \infty$. However, there is no guarantee for Cramér-Rao attainment. If we want an asymptotically efficient estimator, we must instead consider MLE (outside the scope of this question).

Not needed for EC402 but is the only part that I haven't explained above.

"Ragvir, where did Result 1 come from?"

Let $\theta := (\beta', \sigma^2)$. Then, $E_\theta[y_s | X_s, y_s^* > 0] = \int_0^\infty y_s f_{y_s}(y_s | X_s, y_s^* > 0; \theta)\, dy_s$

Notice that $f_{y_s}(y_s | X_s, y_s^* > 0; \theta) = \dfrac{(1/\sigma)\phi((y_s - x_s'\beta)/\sigma)}{\Phi(x_s'\beta/\sigma)}$, a "hazard" function.

$$\therefore E_\theta[y_s | X_s, y_s^* > 0] = \Phi^{-1}(x_s'\beta/\sigma) \int_0^\infty y_s (1/\sigma)\phi((y_s - x_s'\beta)/\sigma)\, dy_s$$

Let $u_s := \phi((y_s - x_s'\beta)/\sigma)$, then

$$-\sigma \frac{du_s}{dy_s} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{y_s - x_s'\beta}{\sigma}\right)^2\right] \cdot \left(\frac{y_s - x_s'\beta}{\sigma}\right), \text{ or}$$

1
$$-\sigma\, du_s = \left[\phi[(y_s - x_s'\beta)/\sigma]\frac{y_s}{\sigma} - \phi\left(\frac{y_s - x_s'\beta}{\sigma}\right)\frac{x_s'\beta}{\sigma}\right] dy_s$$

Clearly if $y_s : 0 \to \infty$, then $u_s : \phi(-x_s'\beta/\sigma) \to 0$

and $-u_s : 0 \to \phi(-x_s'\beta/\sigma)$, or $-u_s : 0 \to \phi\left(\frac{x_s'\beta}{\sigma}\right)$

2

Using substitution 1 and insight 2,

$$\int_0^{\phi(x_s'\beta/\sigma)} \sigma \, du_s = \int_0^\infty \frac{y_s}{\sigma} \phi\left(\frac{y_s - x_s'\beta}{\sigma}\right) dy_s - \frac{x_s'\beta}{\sigma} \int_0^\infty \phi\left(\frac{y_s - x_s'\beta}{\sigma}\right) dy_s$$

$$\sigma \phi(x_s'\beta/\sigma) = \int_0^\infty y_s/\sigma \; \phi\left(\frac{y_s - x_s'\beta}{\sigma}\right) dy_s - \frac{x_s'\beta}{\sigma} \int_0^\infty \left[1 - \Phi\left(\frac{-x_s'\beta}{\sigma}\right)\right] dy_s$$

$$= \int_0^\infty \frac{y_s}{\sigma}\, \phi\left(\frac{y_s - x_s'\beta}{\sigma}\right) dy_s - \frac{x_s'\beta}{\sigma}\left[\Phi\left(x_s'\beta/\sigma\right)\sigma\right]$$

It follows immediately that :

$$E_\theta\left[y_s \mid x_s, y_s^* > 0\right] = x_s'\beta + \sigma\, \frac{\phi(x_s'\beta)}{\Phi(x_s'\beta/\sigma)}$$

which Vassilis calls "Result 1".

I REPEAT : YOU DO NOT NEED TO KNOW THIS DERIVATION OF RESULT 1. IT IS HERE FOR CURIOUS STUDENTS (WHO WERE PLANNING TO ASK ME ABOUT RESULT 1) ONLY.

As we leave NLLS, can you now see the intuition for why OLS cannot work in this case? Let's look at our model again together...

$$E_\theta(y_s | x_s, y_s > 0) = x_s'\beta + \sigma \frac{\phi(x_s'\beta | \sigma)}{\Phi(x_s'\beta | \sigma)}$$

The conditional expectation of y_s given $y_s > 0$ (and x_s) does not depend only on $x_s'\beta$ but also non-linearly on $x_s'\beta$ through the $\phi(\cdot) | \Phi(\cdot)$ function. One could interpret the problem as OVB due to endogeneity arising from a failure to control for the NL term.