London School of Economics
Department of Economics

Vassilis Hajivassiliou
Autumn Term 2024

M.Sc. Econometrics (Ec402)
2024–2025
Answers for Problem Set #5

Answers to Question 1

Consider a linear regression model with data sample (y, X) of size S, that satisfies:

$$A1 : rank(X) = k$$
$$A2 : y = X\beta^{true} + \epsilon^{true}, \qquad E\epsilon^{true} = 0$$
$$A3Rmi : E(\epsilon^{true}|X) = E\epsilon^{true}$$

An investigator wishes to estimate the unknown parameter vector β^{true} and use those estimates to test, in the best possible way, a set of r linear restrictions on this parameter vector specified as: $R\beta^{true} = q$ where R is $r \times k$ and q is $r \times 1$. She is unsure, however, as to which $A4$ assumption characterizes the variance-covariance matrix of the true error vector condtional on the X variables, i.e., what to assume exactly about:

$$varcov(\epsilon^{true}|X) = E(\epsilon^{true}\epsilon^{true\prime}|X)$$

The investigator suspects one of two possible models for $varcov(\epsilon^{true}|X)$ specified as:

Model I	$A4.GM$:	$E(\epsilon^{true}\epsilon^{true\prime}\|X) = \sigma_\epsilon^2 I_S$	GM error vcov
Model II	$A4.\Omega_{ma1}$:	$E(\epsilon^{true}\epsilon^{true\prime}\|X) = \sigma_\epsilon^2 \Omega_{ma1}$	MA1 with $corr(\epsilon_s^{true}, \epsilon_{s-1}^{true}) = \lambda = 0.45$

$$\text{where:} \quad \Omega_{ma1} = \begin{pmatrix} 1 & \lambda & 0 & 0 & \cdots & 0 & 0 \\ & 1 & \lambda & 0 & \ddots & \ddots & 0 \\ & & 1 & \lambda & \ddots & \ddots & \vdots \\ & & & \ddots & \ddots & 0 & 0 \\ & & & & 1 & \lambda & 0 \\ & & & & & 1 & \lambda \\ & & & & & & 1 \end{pmatrix}$$

Consequently, she considers two alternative estimators for the β^{true} coefficient vector:

$$\hat{\beta}_{ols} = (X'X)^{-1}X'y \qquad \hat{\beta}_{igls.ma1} = (X'\Omega_{ma1}^{-1}X)^{-1}X'\Omega_{ma1}^{-1}y$$

1. Suppose that the true model generating the data is Model I. How should the researcher construct two test procedures for testing the r set of linear restrictions, one based on the OLS estimates and another based on the $MA1$ estimates? Which of the two would be preferable?

ANSWER:

Under assumptions A1 and A2, the sampling error vector of $\hat{\beta}_{ols}$ is given by $\hat{\beta}_{ols} - \beta = (X'X)^{-1}X'\epsilon$, while that of $\hat{\beta}_{igls.ma1}$ is given by $(X'\Omega_{ma1}^{-1}X)^{-1}X'\Omega_{ma1}^{-1}\epsilon$. Under A3, both sampling errors have conditional expectation equal to 0, and thus both the OLS and IGLS.MA1 estimators are unbiased.

Under both Models 1 and 2, $\hat{\beta}_{ols}$ will have

$$VCov(\hat{\beta}_{ols}|X) = (X'X)^{-1}X'E(\epsilon\epsilon'|X)X(X'X)^{-1}$$

while the

$$VCov(\hat{\beta}_{igls.ma1}|X) = (X'\Omega_{ma1}^{-1}X)^{-1}X'\Omega_{ma1}^{-1}E(\epsilon\epsilon'|X)\Omega_{ma1}^{-1}X(X'\Omega_{ma1}^{-1}X)^{-1}$$

If the true error model is Model 1, then $E(\epsilon\epsilon'|X) = \sigma^2 I_S$, while under Model 2, then $E(\epsilon\epsilon'|X) = \sigma_\epsilon^2\Omega_{ma1}$.

So under Model 1, the OLS is BLUE by the GM theorem with VCov equal to

$$\sigma^2(X'X)^{-1}$$

while the IGLS.MA1 is *not* BLUE with VCov equal to

$$(X'\Omega_{ma1}^{-1}X)^{-1}X'\Omega_{ma1}^{-1}\Omega_{ma1}^{-1}X(X'\Omega_{ma1}^{-1}X)^{-1}$$

Well-performing test procedures will be obtained by using an estimator and its *correct* VCov matrix under the true error specification.

The *best*-performing (highest power) test procedure will be the one using an estimator that is optimal (e.g., BLUE in this context) and its correct VCov (that will be the smallest possible).

Under error Model 1, the preferred approach will be to use the OLS coefficient estimator and its vcov, since this estimator is BLUE.

2. Now suppose instead that the true model generating the data is Model II. Again the researcher wishes to construct two test procedures based on the two alternative estimators. Should she modify the test procedures from those she gave in part (a)? In which way? Which of the two tests would be preferable in this case?

   ANSWER:

   In this case, the IGLS.MA1 estimator will be BLUE by the GM theorem with the VCov

   $$\sigma_\epsilon^2(X'\Omega_{ma1}^{-1}X)^{-1}X'\Omega_{ma1}^{-1}\Omega_{ma1}\Omega_{ma1}^{-1}X(X'\Omega_{ma1}^{-1}X)^{-1} = \sigma_\epsilon^2(X'\Omega_{ma1}^{-1}X)^{-1}$$

   while OLS will *not* be BLUE, with a VCov given by

   $$\sigma_\epsilon^2(X'X)^{-1}X'\Omega_{ma1}X(X'X)^{-1}$$

   In this case, the preferred test procedure will be the one based on IGLS.MA1 and its correct VCov since this estimator is BLUE in this case. The alternative test procedure based on OLS and its correct VCov matrix given in the last equation, will still be consistent but it will not have the highest possible power since OLS is not BLUE under Model 2.

Answers to Question 2

Consider the linear regression model with k regressors:

$$y_i = x_i'\beta^{true} + \epsilon_i$$

estimated from a (large) cross-section indexed by $i = 1, \ldots, N$. The $N \times k$ regressor matrix X has full column rank k. Every true error disturbance ϵ_i (for all i) is fully statistically independent of the matrix X (i.e., for all regressor variables and all data points).

A researcher wants to estimate the true coefficient vector β^{true} and considers Maximum Likelihood (ML) estimation.

1. The research derives two ML estimators for β^{true} by assuming in turn the following *two alternative error processes*:

   [G] i.i.d. *Gaussian*$(0, v)$ with marginal probability density function (p.d.f.)

   $$f_G(\epsilon_i) = \frac{1}{\sqrt{2\pi v^2}} \exp(-\frac{\epsilon_i^2}{2v^2})$$

   [L] i.i.d. *Logistic*$(0, v)$ with marginal probability density function (p.d.f.)

   $$f_L(\epsilon_i) = \frac{\exp\left(-\frac{\epsilon_i}{v}\right)}{v\left(1 + \exp\left(-\frac{\epsilon_i}{v}\right)\right)^2}$$

   Reminder: The parameter v is the *scale* parameter of each distribution — it is not necessarily equal to the variance of ϵ_s in all cases.

   Denote the three ML estimators she will obtain by $\hat{\beta}_{ml.G}$, $\hat{\beta}_{ml.L}$, and $\hat{\beta}_{ml.C}$ respectively. Set up the log likelihood optimization problem for each of these three estimators. Which, if any, of these formulae have closed-form solutions?

   ANSWER:

   For simplicity, assume that v is known. Then

   $$\begin{aligned} \hat{\beta}_{ml.G} &= \arg\max_b \sum_{i=1}^N \ln f_G(y_i - x_i'b) = \arg\max_b \sum_{i=1}^N \ln \exp(-\frac{(y_i - x_i'b)^2}{2v^2}) \\ &= \arg\min_b \sum_{i=1}^N (y_i - x_i'b)^2 \end{aligned}$$

   $$\begin{aligned} \hat{\beta}_{ml.L} &= \arg\max_b \sum_{i=1}^N \ln f_L(y_i - x_i'b) = \arg\max_b \sum_{i=1}^N \ln \frac{\exp\left(-\frac{y_i - x_i'b}{v}\right)}{v\left(1 + \exp\left(-\frac{y_i - x_i'b}{v}\right)\right)^2} \\ &= \arg\max_b \sum_{i=1}^N \left\{ \left(-\frac{y_i - x_i'b}{v}\right) - \ln v\left(1 + \exp\left(-\frac{y_i - x_i'b}{v}\right)\right)^2 \right\} \end{aligned}$$

   Only the formula for ml.G has a closed form solution — the ML.G estimator coincides with OLS. The other one (ml.L) will require iterative solution on a computer.

2. Are there conditions that would make one or both of these estimators *Best Linear Unbiased* (*BLUE*)? *BUE*?

   ANSWER:

   Since ml.G is the same as OLS, it will be BLUE under A1+A2+A3Rmi+A4GM; and BUE (and also Best CUAN) if we add +A5G.

   The other one (ml.L) is a non-linear estimator, and hence can be neither BLUE nor BUE. (Because of their non-linearity, the presumption is that they will be biased in finite samples.)

1 Answers to Question 3

We are given $X'X = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $X'y = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $y'y = \frac{4}{3}$.

$\widehat{Ey}_{12} = \hat{y}_{12} = \frac{1}{3}5 + \frac{1}{3}(-2) = 1$

$\widehat{Ey}_{13} = \hat{y}_{13} = \frac{1}{3}3 + \frac{1}{3}(-7) = -4/3$

$\widehat{var}(\widehat{Ey}_{12}) = \frac{1}{81}\begin{bmatrix} 5 & -2 \end{bmatrix}\begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} 5 \\ -2 \end{bmatrix} = 156/81 = 1.92$

$\widehat{var}(\widehat{Ey}_{13}) = \begin{bmatrix} 3 & -7 \end{bmatrix}\begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} 3 \\ -7 \end{bmatrix} = 316/81 = 3.901$

$\widehat{var}(\widehat{y}_{12}) = s^2 + \widehat{var}(\widehat{Ey}_{12}) = (6 + 156)/81 = 2$

$\widehat{var}(\widehat{y}_{13}) = s^2 + \widehat{var}(\widehat{Ey}_{12}) = (6 + 316)/81 = 3.975$

Since the question asks for an 80% prediction interval, we need the critical values from the $t(9)$ distribution with 10% in each tail, i.e., ± 1.383.

1. 80% prediction interval for $\hat{y}_{12}$: $1 \pm 1.383\sqrt{2}$
   80% prediction interval for $\hat{y}_{13}$: $-4/3 \pm 1.383\sqrt{3.975}$

2. 80% prediction interval for $\widehat{Ey}_{12}$: $1 \pm 1.383\sqrt{1.92}$
   80% prediction interval for $\widehat{Ey}_{13}$: $-4/3 \pm 1.383\sqrt{3.901}$.

3. The prediction intervals for $\hat{y}_{12}$ and $\hat{y}_{13}$ are wider than the ones for the expected values of the y's because extra prediction variance is introduced as ϵ_{12} and ϵ_{13} are also being predicted (to be 0) in that case.

2 Answers to Question 4

1. Consider the linear regression model:

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + \epsilon_t, \quad t = 1, \cdots, T,$$

with $\rho(X'X) = 4$ and $\epsilon|X \sim N(0, \sigma_\epsilon^2 I_T)$. We are interested in the following hypotheses:

(a) H_0: $\beta_2 - 3\beta_3 = 4$, $\beta_1 = 2\beta_4$

(b) H_0': $\beta_1 = 1$, $\beta_2 = 3\beta_4 - 1$, $\beta_3 = 0$

(c) H_0'': $\beta_2 = \beta_3 = 2\beta_4$

i. Explain how you would use a statistic of the form $W = (R\hat{\beta} - q)'(R\hat{V}(\hat{\beta})R'^{-1}(R\hat{\beta} - q)/r$ to carry out tests of the 3 null hypotheses above. Be explicit about the quantities R, q, and r, and the degrees of freedom involved.

ANSWER:

For H0: $r = 2$, $R = \begin{bmatrix} 0 & 1 & -3 & 0 \\ 1 & -2 & 0 & 0 \end{bmatrix}$, $q = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$

For H0': $r = 3$, $R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, $q = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$

For H0": $r = 2$, $R_a = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}$, $q = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

OR, equivalently, $R_b = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}$ since $\beta_2 = \beta_3 = 2\beta_4$ can be written as two separate restrictions:

$\beta_2 = \beta_3$ and $\beta_2 = 2\beta_4$ OR $\beta_2 = \beta_3$ and $\beta_3 = 2\beta_4$