MSc/PGC/MRes Econometrics, Tutorial Class Exercises.

Ron Smith & Walter Beckert

Autumn 2020

Department of Economics, Mathematics and Statistics,
Birkbeck, University of London

MSc Economics, Financial Economics, Finance, Financial Risk Management, Financial Engineering.
PGCert Econometrics

These exercises should be attempted before coming to class. Unless you try to do them yourself you will not know what problems you are having.

1 Week 2

Consider the Linear Regression Model

$$y = X\beta + u$$

where y is a $T \times 1$ vector of observations on a dependent variable; X a $T \times k$ rank k matrix of observations on non-stochastic exogenous variables; u a $T \times 1$ vector of unobserved disturbances with $E(u) = 0$, $E(uu') = \sigma^2 I$, and β a $k \times 1$ vector of unknown coefficients.

The Least Squares estimator of β is $\widehat{\beta} = (X'X)^{-1}X'y$.

(a) The Gauss-Markov theorem is that $\widehat{\beta}$ is the Best Linear Unbiased Estimator (BLUE) of β. Prove the Gauss-Markov theorem and explain what role each of the assumptions play in the proof. How would the derivation change if the exogenous variables X were stochastic, but distributed independently of the errors.

(b) Define $P_X = X(X'X)^{-1}X'$ and $M = I_T - P_X$. Show: (i) $MM = M$, (ii) $MP_X = 0$. (iii) $\widehat{u} = My = Mu$.

(c) Show that $E(\widehat{u}'\widehat{u}) = (T-k)\sigma^2$.

(d) Suppose that $E(uu') = \sigma^2\Omega$. (i) What is $E(\widehat{\beta} - E(\widehat{\beta}))((\widehat{\beta} - E(\widehat{\beta}))'$? (ii) Derive the estimator $\widetilde{\beta}$ that makes $X'\Omega^{-1}\widetilde{u} = 0$, where $\widetilde{u} = y - X\widetilde{\beta}$.

2 Week 3

Suppose for $t = 1, 2, ..., T$

$$Y_t = \alpha + u_t$$

where the observations are independent and the distribution of Y_t is given by

$$f(Y_t) = (2\pi\sigma^2)^{-1/2} \exp -\frac{1}{2}(\frac{Y_t - \alpha}{\sigma})^2$$

(a) What is the log-likelihood function and the score vector? Derive the maximum likelihood estimators of $\theta = (\alpha, \sigma^2)$

(b) Derive the information matrix and the asymptotic variance-covariance matrix.

(c) How would you estimate the standard error of the maximum likelihood estimator of α?

(d) Compare your derivations to the matrix form for the linear regression model in the notes.

3 Week 4

(a) In the Linear Regression Model in the week 2 exercise, suppose that the disturbances are also normally distributed and there are k prior restrictions of the form $\beta - q = 0$, where q is a known vector of order $k \times 1$. Derive a test statistic to test these restrictions. Explain how you would calculate the restricted and unrestricted sums of squares to carry out the test.

(b) The following equations were estimated on 24 observations 1918-1941, where D_t is dividends and E_t is earnings in year t. Standard errors are given in parentheses, SSR is sum of squared residuals, MLL is maximised log likelihood.

$D_t =$	0.59	$+0.40E_t$			$SSR = 2.1849$
	(0.20)	(0.10)			$MLL = -5.297$
$D_t =$	-0.14	$+0.32E_t$	$-0.10E_{t-1}$	$+0.70D_{t-1}$	$SSR = 0.84821$
	(0.17)	(0.08)	(0.10)	(0.14)	$MLL = 6.0576$

Test the hypothesis that the coefficient of earnings in the first equation (i) equals one (ii) equals zero.

Test the hypotheses that the coefficients of lagged earnings and dividends in the second equation equal zero, (i) individually (ii) jointly. For the joint test use both F and LR tests.

Suppose the coefficients in the second equation are labelled $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)$. Write the restrictions that the coefficients of lagged earnings and dividends equal zero in the form $R\beta - q = 0$.

4 Week 5

1. Explain what effect the following 'problems' have on the properties of the least squares estimates of the coefficients and their standard errors. How would you detect whether each problem was present:

(a) Heteroskedasticity.
(b) Serial correlation.
(d) Non-normality.
(e) Non-linearity.
(f) Exact multicollinearity.

2. Explain the following quote, from Angrist and Pischke (2009) p223. Do you agree with it? 'We prefer fixing OLS standard errors to GLS. GLS requires even stronger assumptions than OLS, and the resulting asymptotic efficiency gain is likely to be modest, while finite sample properties may be worse.'

5 Week 6

Reading Week, Computer classes.

6 Week 7

Consider the following models estimated over a sample $t = 1, 2, ..., T$. In each case ε_t is white noise and ρ and μ are less than one in absolute value.

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t \tag{1}$$

$$y_t = \alpha + y_{t-1} + \varepsilon_t \tag{2}$$

$$y_t = \alpha + \varepsilon_t + \mu\varepsilon_{t-1} \tag{3}$$

$$\Delta y_t = \alpha + \rho\Delta y_{t-1} + \varepsilon_t + \mu\varepsilon_{t-1} \tag{4}$$

(a) Suppose you had estimates of the parameters, how would you forecast y_{T+1} and y_{T+2} in each case, given data up to y_T?

(b) In cases (1) and (2) substitute back to remove the y_{t-i}.

(c) For cases (1) and (3) what is the expected value of y_t?

(d) For (1) and (3) derive the variance, covariances and autocorrelations of the series.

(e) For which of the models is y_t I(1)?

7 Week 8

The Augmented Dickey Fuller Tests for non-stationarity uses a regression of the form:

$$\Delta y_t = \alpha + \beta y_{t-1} + \gamma t + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \varepsilon_t$$

(a) What is the null hypothesis tested; what is the test statistic and what is its 95% critical value?

(b) Suppose all the $\delta_i = 0$. Substitute back to express y_t in terms of ε_{t-i}, t and y_0. Compare the cases $\beta < 0$ and $\beta = 0$.

(c) What is the rationale behind the $\sum_{i=1}^{p} \delta_i \Delta y_{t-i}$ term.

(d) using the Shiller data test whether the log price-earnings ratio, $\log(NSP/NE)$, has a unit root (i) over 1950-1999, (ii) over the whole sample; using intercept and no trend and using the AIC to choose p.

8 Week 9

Consider the general model

$$d_t = \alpha_0 + \alpha_1 d_{t-1} + \beta_0 e_t + \beta_1 e_{t-1} + \gamma_0 p_t + \gamma_1 p_{t-1} + u_t$$

where d_t is log nominal dividends, e_t is log nominal earnings and p_t is the log of the producer price index.

(a) How would you calculate the long-run elasticity of dividends to prices and earnings, θ_i in:

$$d_t^* = \theta_0 + \theta_1 e_t + \theta_2 p_t.$$

.

(b) How would you estimate the model subject to the restrictions that the long-run elasticity to earnings is unity and to prices zero.

(c) Suppose it was believed that the appropriate model was

$$d_t = \alpha + \beta e_t + \gamma p_t + v_t;\ v_t = \rho v_{t-1} + \varepsilon_t,$$

where ρ is less than one in absolute value. Show that this is a restricted form of the general model and derive the two restrictions.

(d) Suppose it was believed that the adjustment to the long-run relationship was given by:

$$\Delta d_t = \lambda_1 \Delta d_t^* + \lambda_2 (d_{t-1}^* - d_{t-1}) + u_t$$

what restriction does this impose on the general model.

(e) Using the Shiller16 data, estimate the models over the period 1950-1990 and test the three sets of restrictions.

9 Week 10.

Consider the VAR

$$y_t = A_0 + A_1 y_{t-1} + \varepsilon_t;\ t = 1, 2, ..., T$$

where $y_t = (y_{1t}, y_{2t})'$, A_0 is a $2x1$ vector, $(a_o, a_1)'$,

$$A_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \varepsilon_t \tilde{}N \left(\begin{matrix} 0 \\ 0 \end{matrix}, \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix} \right)$$

(a) Write down each equation of the VAR and explain how you would estimate the coefficients and covariance matrix.

(b) Given estimates of the parameters, how would you forecast y_{T+1} and y_{T+2}?

(c) What condition is required for y_{1t} to be Granger non-causal with respect to y_{2t}?

(d) Write the VAR as

$$\Delta y_t = A_0 + \Pi y_{t-1} + \varepsilon_t;\ t = 1, 2, ..., T.$$

Explain the relation between A_1 and Π.

(e) What are the implications for Π if y_{it} are (i) I(0); (ii) I(1) and cointegrated; (iii) I(1) and not cointegrated? What restrictions does case (iii) put on A_1?

(f). Suppose the y_{it} are I(1) with cointegrating relationship $z_t = y_{1t} - \beta y_{2t}$. Write out the restricted system and explain the restrictions this imposes on Π. Show that it has rank one.

(g) Derive the parameters of the ARDL(1,1) model:

$$y_{1t} = \alpha_0 + \beta_0 y_{2t} + \beta_1 y_{2t-1} + \alpha_1 y_{1t-1} + u_t$$

from the VAR. Hint note that $E(\varepsilon_{1t} \mid \varepsilon_{2t}) = (\omega_{12}/\omega_{22})\varepsilon_{2t}$ and use this in the first equation of the VAR substituting for ε_{2t}.

10 Week 11

Consider the model

$$y = \mathbf{X}\beta + u, \qquad u \sim N(0, \sigma^2 I),$$

where y is a $T \times 1$ vector of observations on a dependent variable, $\mathbf{X}$ is a $T \times k$ full-rank matrix of observations on potentially endogenous variables, β is a $k \times 1$ vector of unknown parameters, and u is a $T \times 1$ vector of unobserved disturbances. Suppose that there also exists a $T \times m$ matrix of instruments, $\mathbf{W}$, such that $E(\mathbf{W}'\mathbf{X}) \neq 0$ and $E(\mathbf{W}'u) = 0$. $\mathbf{W}$ will contain elements of $\mathbf{X}$ that are exogenous.

(a) Explain the conditions for β to be just-identified and to be over-identified.

(b) Suppose β was just-identified; what is the instrumental variable estimator, $\widetilde{\beta}$, and how would you estimate its variance–covariance matrix?

(c) Suppose β was over-identified; what is the generalised instrumental variable estimator and how would you estimate its variance–covariance matrix?

(d) Suppose β was over-identified; how would you (i) test the over-identifying restrictions and (ii) determine whether the instruments were weak?

(e) Consider the simple Keynesian system, for income, consumption, investment and government expenditure for $t = 1, 2, ..., T$

$$\begin{aligned} Y_t &= C_t + I_t + G_t \\ C_t &= \alpha + \beta Y_t + u_t. \end{aligned}$$

Is the consumption function just identified or overidentified? How would you estimate the consumption function and test any overidentifying restrictions.