Econ 7010 - Final - Fall 2023

Ryan T. Godwin

The exam is 180 minutes long, and has 100 marks. You may not use any outside materials, only writing implements. Write your answers in the booklet provided.

Short Answer - Answer 10 out of 12 questions. Only the first 10 questions will be marked. 30% total, each question worth 3%.

1. Prove that the least squares estimator is unbiased and consistent, stating any assumptions that you use.

   Unbiased: see equation 5.1 on pg. 48. Consistent: see the bottom of pg. 64 and top of pg. 65.

   ---

2. Explain the theoretical and practical reasons for deriving the variance-covariance matrix of $\boldsymbol{b}$: $V(\boldsymbol{b}) = \sigma^2(X'X)^{-1}$. (What results did we derive using this matrix, and what is this matrix used for in practice?)

   This matrix is needed in the Gauss-Markov theorem to prove the efficiency of LS. It is also needs to be estimated (by replacing σ^2 with s^2) in order to calculate standard errors, confidence intervals, t, F, z, and Wald tests, and to perform hypothesis testing in general.

   ---

3. Which do you think is the most important of the LS assumptions (A.1 - A.6), and why?

   A.5 is likely the most important assumption; it is needed for both unbiasedness and consistency, and is impossible to test. You could argue for a different assumption depending on context however.

   ---

4. Show how to estimate a Cobb-Douglas production function using LS.

   See the top of pg. 16.

   ---

5. Why does R^2 always increase when a regressor is added tot he model, and how does $\bar{R}^2$ fix the problem?

   See pg. 41.

   ---

6. Explain how a critical value, a confidence interval **or** a p-value is used for hypothesis testing.

Explain the decision rules here. For example: a critical value is the maximum value for the test statistic before the null hypothesis is rejected, given some significance level.

---

7. Show how to write the null hypothesis $H_0 : \beta_1 = 0, \beta_2 = \beta_3$ in terms of R and $\boldsymbol{q}$, where $k = 3$.

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

$$q = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

---

8. Explain how to do White's test for heteroskedasticity.

See pg. 105.

---

9. What is the difference between the GLS and FGLS estimator?

See section 11.5 on pg. 110.

---

10. What is autocorrelation, and why is it common in time series data?

See the first paragraph of section 12.2 on pg. 112.

---

11. Derive the variance of ϵ_t, where ϵ_t follows an AR(1) process. What condition is required for the variance to be finite?

See the middle of pg. 116.

---

12. Explain what a spurious regression is, in the context of two random walks.

Se section 12.4 on pg. 117.

Long Answer - Answer 4 out of 5 questions. Only the first 5 questions will be marked. 70% total, each question worth 17.5%, each part worth 3.5%.

1. Asymptotics.

   a) Under A.6 (Normality), $\mathrm{E}(s^2) = \sigma^2$ and $\mathrm{var}(s^2) = 2\sigma^4/(n-k)$. Show that s^2 is mean-square consistent (strongly consistent).

   See the bottom theorem on pg. 65.

   ---

   b) Using the result that s^2 is consistent, and Slutsky's theorem, show that $\hat{\sigma}^2$ is at least *weakly* consistent.

   Several ways to answer this.

   (i) One estimator has an $n-k$ in the denominator, and the other has just n in the denominator. As $n \to \infty$ there is no difference between the two. So if one is consistent, so is the other.

   (ii)
$$s^2 = \frac{\boldsymbol{e}'\boldsymbol{e}}{n-k}$$
   and
$$\hat{\sigma}^2 = \frac{\boldsymbol{e}'\boldsymbol{e}}{n} = s^2\frac{n-k}{n}$$
   As $n \to \infty$, $\hat{\sigma}^2 \to s^2$, so if one is consistent both are consistent.

   (iii) Use Slutsky's theorem, which says that if we know the plim of $\hat{\theta}$, we know the plim of any function of $\hat{\theta}$.

$$\mathrm{plim}(s^2) = \sigma^2$$
$$\hat{\sigma}^2 = s^2\frac{n-k}{n}$$
$$\mathrm{plim}(\hat{\sigma}^2) = \mathrm{plim}(s^2) \times \lim_{n\to\infty} \frac{n-k}{n} = \sigma^2$$

   ---

   c) Prove that the LS estimator is consistent, stating your assumptions.

   See pg. 64 - 65.

   ---

   d) Under what circumstances can the LS estimator be inconsistent?

   Some examples that we have seen in class: when there is an omitted variable that is correlated with X, when a multiplicative model is log-linearized and the error term is heteroskedastic, and when a lagged dependent variable is included as a regressor when the error term follows an AR process.

   ---

e) Why do we need to consider the asymptotic distribution of $\sqrt{n}\,(\boldsymbol{b}-\boldsymbol{\beta})$ instead of the asymptotic distribution of just $\boldsymbol{b}$?

See section 7.3.

---

2. Instrumental variables.

a) What problem is IV estimation trying to fix?

The inconsistency of the LS estimator due to a missing variable (or due to simultaneity).

---

b) Derive the IV estimator (either the simple or over-identified), any way you like.

See the first half of section 8.3.

---

c) What properties must an instrument have, in order for it to be "valid"?

See "relevance" and the "exclusion restriction" at the top of section 8.2.

---

d) Briefly explain the Hausman test.

See section 8.4.1

---

e) What is a "weak" instrument?

See section 8.4.3

---

3. Multiple hypothesis testing and restricted least squares.

a) When would you use an F-test instead of a t-test?

Any time you are testing a null hypothesis involving more than one coefficient, you need to use an F-test rather than a t-test.

---

b) How can you perfrom an F-test by estimating two different models?

The null hypothesis suggests a "restricted" version of the model under the alternative hypothesis. That is, there are two models, the restricted model under the null, and the unrestricted under the alternative. By obtaining the R-squared from these two models, the F-test statistic may be calculated.

---

c) If A.6 (Normality of the error terms) is violated, what happens to the F-test? What test could you use instead?

If A.6 is violated, the F-stat no longer necessarily follows the F-distribution. However, by modifying the F-stat slightly (multiply it by the number of restrictions being tested, J), we get the Wald statistic, which is chi-square distributed regardless of the distribution of ϵ.

---

d) Explain how you would use a dummy variable that fully interacts with every X variable, and an F-test, to test for differences between two groups.

See section 9.6.

---

e) Using the RLS estimator, explain the potential benefits and costs of imposing restrictions on a model.

The RLS estimator is in equation 9.6 on pg. 84. By taking the expected value and plim of this estimator we see that RLS is biased and inconsistent unless the restrictions are true. This is the potential cost of imposing restrictions. By taking the variance of the RLS estimator, we see that it is less than that of LS, regardless of whether the restrictions are true. This is the potential benefit of using RLS.

---

4. Non-linear effects and models.

a) Other than estimating a non-linear model, what is a way that we can deal with non-linear relationships between variables?

We can attempt to linearize the model. We can use polynomials, logs, and interactions to approximate the non-linear relationship between variables.

---

b) What is the non-linear least squares estimator? (How is the estimator derived? How is the estimator mathematically different from the LS estimator?)

See equation 10.5 and the two paragraphs below it.

---

c) Derive the Newton-Raphson algorithm graphically.

See figure 10.4.

---

d) Explain a situation where the Newton-Raphson algorithm may fail to converge.

See some of the bullet points at the top of pg. 96.

---

e) Explain why we might want to estimate the following gravity model using NLS:

$$T_{ij} = \alpha_0 Y_i^{\alpha_1} Y_j^{\alpha_2} D_{ij}^{\alpha_3} \eta_{ij}$$

See section 10.5.

---

5. Heteroskedasticity.

a) What are the mathematical definitions of homoskedasticity and heteroskedasticity?

See the intro to chapter 11.

---

b) What are the consequences of heteroskedasticity?

See section 11.1.

---

c) What are some solutions to heteroskedasticity? (How can you deal with it?)

You can fix the inconsistency of the standard errors using a robust estimator (such as White's). You can fix both the standard errors and the inefficiency of LS using the GLS or FGLS estimator.

---

d) Suppose that you are dealing with "clustered" or grouped data (the observations have been averaged over groups). There are two different group sizes, and you know which observation belongs to each group. The group sizes are 100 and 400. What is the GLS estimator for this data?

The GLS estimator is given by equation 11.8 on pg. 109, and in this case Ω is:

$$\Omega = \begin{bmatrix} \frac{1}{100} & & & & & \\ & \ddots & & & & \\ & & \frac{1}{100} & & & \\ & & & \frac{1}{400} & & \\ & & & & \ddots & \\ & & & & & \frac{1}{400} \end{bmatrix}$$

and the P matrix will be:

$$P = \begin{bmatrix} 10 & & & & & \\ & \ddots & & & & \\ & & 10 & & & \\ & & & 20 & & \\ & & & & \ddots & \\ & & & & & 20 \end{bmatrix}$$

In GLS, the data is weighted. In this case, each observation from the $n = 100$ group gets multiplied by 10, and from the $n = 400$ group gets multiplied by 20. (Or just multiply the second group by 2, since all that matters is that weights are *proportional* to the variances.

---

e) How does the GLS estimator above solve the problem of heteroskedasticity?

See equation 11.7 on pg. 108. Since the transformed model has a homoskedastic error term, the usual covariance matrix $\sigma^2(X'X)^{-1}$ can be applied to estimate standard errors, t-stats, etc.

END.