Economics 468 – Midterm Exam

1. Consider the linear regression model

$$y_t = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + u_t.$$

Explain how you could estimate this model subject to the restriction that $\beta_2 + \beta_3 = 1$ by running a regression that imposes the restriction. Also, explain how you could estimate the unrestricted model in such a way that the value of one of the coefficients would be zero if the restriction held exactly for your data.

If $\beta_2 + \beta_3 = 1$, that is, $\beta_3 = 1 - \beta_2$, the model becomes

$$y_t - x_{t3} = \beta_1 + \beta_2(x_{t2} - x_{t3}) + u_t. \tag{S.01}$$

Running this regression imposes the restriction.

If we keep $x_{t2} - x_{t3}$ as a regressor, and then include another regressor, which could be either x_{t2} or x_{t3}, we get (with x_{t2})

$$y_t - x_{t3} = \beta_1 + \beta_2(x_{t2} - x_{t3}) + \gamma_2 x_{t2} + u_t. \tag{S.02}$$

The regressors in this model clearly span the same space as do the constant, x_{t2} and x_{t3}, and so both models contain exactly the same DGPs. Comparing (S.01) and (S.02) shows that, if (S.01) is correctly specified, then $\gamma_2 = 0$.

2. The class of estimators considered by the Gauss-Markov Theorem for the linear regression $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}$ can be written as $\tilde{\boldsymbol{\beta}} = \boldsymbol{Ay}$, for some exogenous matrix $\boldsymbol{A}$ with $\boldsymbol{AX} = \mathbf{I}$. Show that any estimator of this form is unbiased.

If the true DGP is $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}$, then

$$\tilde{\boldsymbol{\beta}} = \boldsymbol{A}(\boldsymbol{X\beta} + \boldsymbol{u}) = \boldsymbol{\beta} + \boldsymbol{Au}.$$

Thus the estimation error $\tilde{\boldsymbol{\beta}} - \beta = \boldsymbol{Au}$, of which the expectation is zero, since $\boldsymbol{A}$ is exogenous.

Show that this class of estimators is in fact identical to the class of estimators of the form

$$\tilde{\boldsymbol{\beta}} = (\boldsymbol{W}^\top\boldsymbol{X})^{-1}\boldsymbol{W}^\top\boldsymbol{y}, \tag{S.03}$$

where $\boldsymbol{W}$ is a matrix of exogenous variables such that $\boldsymbol{W}^\top\boldsymbol{X}$ is nonsingular.

It is clear that any estimator of the form (S.03) is also of the form $\boldsymbol{Ay}$, with $\boldsymbol{A} = (\boldsymbol{W}^\top\boldsymbol{X})^{-1}\boldsymbol{W}^\top$, which satisfies the requirement that $\boldsymbol{AX} = \mathbf{I}$.

For the other direction, we can simply let $\boldsymbol{W} = \boldsymbol{A}^\top$, for, in this case, $\boldsymbol{W}^\top\boldsymbol{X} = \boldsymbol{AX} = \mathbf{I}$, so that $(\boldsymbol{W}^\top\boldsymbol{X})^{-1}\boldsymbol{W}^\top\boldsymbol{y} = \boldsymbol{W}^\top\boldsymbol{y} = \boldsymbol{Ay}$.

If $\tilde{\boldsymbol{\beta}}$ is $k \times 1$, and $\boldsymbol{w}$ is a non-random or exogenous $k \times 1$ vector, what is the variance of $\boldsymbol{w}^\top\tilde{\boldsymbol{\beta}}$ as a function of the covariance matrix of $\tilde{\boldsymbol{\beta}}$?

Since we are supposing that $\tilde{\boldsymbol{\beta}}$ is unbiased, the variance of $\boldsymbol{w}^\top\boldsymbol{\beta}$ is

$$\mathrm{E}\big(\boldsymbol{w}^\top(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})^\top\boldsymbol{w}\big).$$

Let $\boldsymbol{V}$ denote the covariance matrix of $\tilde{\boldsymbol{\beta}}$. Then, since $\boldsymbol{w}$ is exogenous, the expression above for the variance is just $\boldsymbol{w}^\top\boldsymbol{V}\boldsymbol{w}$.

3. Let the scalar random variable Y_n have cumulative distribution function (CDF)

$$F_n(x) = \begin{cases} 0 & \text{for } x < 0, \\ nx & \text{for } 0 \le x \le 1/n, \\ 1 & \text{for } x > 1/n. \end{cases}$$

Y_n is said to follow the uniform distribution $\mathrm{U}(0, 1/n)$, since its density is constant and equal to n on the interval $[0, 1/n]$, and zero elsewhere. Show that the sequence $\{Y_n\}$ converges in distribution. What is the limiting CDF F_∞? Show that F_∞ has a point of discontinuity at 0, and that $\lim_{n\to\infty} F_n(0) \ne F_\infty(0)$.

For $x < 0$, it is immediate that $\lim_{n\to\infty} F_n(x) = 0$, and so $F_\infty(x) = 0$. For $x > 0$, choose N such that $N > 1/x$, so that $x > 1/N$. Then, for all $n > N$, $F_n(x) = 1$, and so $\lim_{n\to\infty} F_n(x) = 1$, implying that $F_\infty(x) = 1$. For $x = 0$, $F_n(0) = 0$, and so $\lim_{n\to\infty} F_n(0) = 0$. But, since a CDF must be continuous to the right, the limiting CDF F_∞ satisfies the property

$$F_\infty(0) = \lim_{x\downarrow 0} F_\infty(x) = 1.$$

It follows that F_∞ is discontinuous at $x = 0$, and that $\lim_{n\to\infty} F_n(0) \ne F_\infty(0)$. In fact, F_∞ is the CDF of a degenerate random variable that is equal to zero.

4. Another exercise dealing with the consumption function. The time series c_t is the log of consumption expenditures, and y_t is the log of disposable income. The model estimated is:

$$\Delta c_t = \beta_1 + \beta_2\Delta y_t + \beta_3\Delta y_{t-1} + \beta_4\Delta y_{t-2} + u_t, \tag{S.04}$$

where Δ is the first-difference operator, so that, for instance, $\Delta c_t = c_t - c_{t-1}$. Here are the results of the estimation.

(output omitted)

Let $\gamma = \sum_{i=2}^{4} \beta_i$. Compute the OLS estimate, $\hat{\gamma}$ of the parameter γ. Compute its variance and its standard error.

For $\hat{\gamma}$ it is enough to add up the OLS estimates of the β_i, $i = 2, 3, 4$. We find that $\hat{\gamma} = 0.677242$.

For the standard error, we can use the result of question 2. The covariance matrix $\boldsymbol{V}$ is shown in the output, and we want $\boldsymbol{w} = [0 \;\; 1 \;\; 1 \;\; 1]^\top$, so as to get the variance of $\hat{\gamma}$. On taking the square root, we find that the standard error is 0.076478.

An econometrician decided to run another regression with the same data, as follows:

$$\Delta c_t = \theta_1 + \theta_2 \Delta y_t + \theta_3(\Delta y_{t-1} - \Delta y_t) + \theta_4(\Delta y_{t-2} - \Delta y_t) + u_t. \tag{S.05}$$

Express the coefficients θ_i, $i = 1, 2, 3, 4$, as functions of the β_i, $i = 1, 2, 3, 4$. For an econometrician interested in the parameter γ, why is this second regression useful? In particular, explain why the output from this second regression provides the estimate $\hat{\gamma}$, and its standard error.

It is clear that the regressors in (S.05) span the same space as those in (S.04). Thus the relation between the β_i and the θ_i follows from

$$\beta_1 + \beta_2 \Delta y_t + \beta_3 \Delta y_{t-1} + \beta_4 \Delta y_{t-2} = \theta_1 + \theta_2 \Delta y_t + \theta_3(\Delta y_{t-1} - \Delta y_t) + \theta_4(\Delta y_{t-2} - \Delta y_t).$$

Given the linear independence of the constant, Δy_t, and its lags, this implies that

$$\theta_1 = \beta_1, \quad \theta_3 = \beta_3, \quad \theta_4 = \beta_4, \quad \text{and} \quad \beta_2 = \theta_2 - \theta_3 - \theta_4.$$

From this, we see that $\theta_2 = \beta_2 + \beta_3 + \beta_4 = \gamma$. This explains the econometrician's interest in the second regression, because one can read off the estimate $\hat{\gamma}$ as $\hat{\theta}_2$, as also its standard error.

5. A theorem states that the covariance matrix $\mathrm{Var}(\boldsymbol{b})$ of any random k–vector $\boldsymbol{b}$ is positive semidefinite. Prove this fact by considering arbitrary linear combinations $\boldsymbol{w}^\top \boldsymbol{b}$ of the components of $\boldsymbol{b}$ with nonrandom $\boldsymbol{w}$. If $\mathrm{Var}(\boldsymbol{b})$ is positive semidefinite without being positive definite, what can you say about $\boldsymbol{b}$?

We saw in question 2 that

$$\mathrm{Var}(\boldsymbol{w}^\top \boldsymbol{b}) = \boldsymbol{w}^\top \mathrm{Var}(\boldsymbol{b})\, \boldsymbol{w}. \tag{S.06}$$

Since variances cannot be negative, the expression on the right-hand side above must be greater than or equal to zero. But, if this is true for all $\boldsymbol{w}$, then $\mathrm{Var}(\boldsymbol{b})$ must be positive semidefinite, since this is the definition of a positive semidefinite matrix.

If $\mathrm{Var}(\boldsymbol{b})$ is positive semidefinite but not positive definite, there must exist at least one non-trivial linear combination of the elements of $\boldsymbol{b}$, say $\boldsymbol{w}_0{}^\top \boldsymbol{b}$, that has variance 0. For example, the elements of $\boldsymbol{b}$ might sum to a constant, in which case $\boldsymbol{w}_0$ would be a vector of 1s.

For any pair of scalar random variables, b_1 and b_2, show, by using the fact that the covariance matrix of $\boldsymbol{b} \equiv [b_1 \,\vdots\, b_2]$ is positive semidefinite, that

$$\big(\mathrm{cov}(b_1, b_2)\big)^2 \le \mathrm{Var}(b_1)\,\mathrm{Var}(b_2). \tag{S.07}$$

Use this result to show that the correlation of b_1 and b_2 lies between -1 and 1.

The easiest way to proceed is to use the fact that the determinant of a positive semidefinite matrix is nonnegative. The determinant of

$$\operatorname{Var}(\boldsymbol{b}) = \begin{bmatrix} \operatorname{Var}(b_1) & \operatorname{cov}(b_1, b_2) \\ \operatorname{cov}(b_1, b_2) & \operatorname{Var}(b_2) \end{bmatrix}$$

is just $\operatorname{Var}(b_1)\operatorname{Var}(b_2) - \big(\operatorname{cov}(b_1, b_2)\big)^2$. Since this must be nonnegative, the inequality (S.07) follows immediately.

By the definition of a correlation, we have

$$\rho(b_1, b_2) \equiv \frac{\operatorname{cov}(b_1, b_2)}{\big(\operatorname{Var}(b_1)\operatorname{Var}(b_2)\big)^{1/2}}.$$

It follows from (S.07) that the square of this is no greater than 1. Thus $\rho(b_1, b_2)$ itself must be less than 1 in absolute value.

If one does not wish to use determinants, (S.07) can be shown directly by the following slightly tricky argument. Let $s_1^{-2} = \operatorname{Var}(b_1)$ and $s_2^{-2} = \operatorname{Var}(b_2)$. Then

$$\begin{aligned} &[\, s_1 \quad -s_2 \,] \begin{bmatrix} \operatorname{Var}(b_1) & \operatorname{cov}(b_1, b_2) \\ \operatorname{cov}(b_1, b_2) & \operatorname{Var}(b_2) \end{bmatrix} \begin{bmatrix} s_1 \\ -s_2 \end{bmatrix} = \\ &\quad s_1^2 \operatorname{Var}(b_1) + s_2^2 \operatorname{Var}(b_2) - 2\, s_1 s_2 \operatorname{cov}(b_1, b_2) = 2 - \frac{2\operatorname{cov}(b_1, b_2)}{\big(\operatorname{Var}(b_1)\operatorname{Var}(b_2)\big)^{1/2}}. \end{aligned}$$

Since this must be nonnegative, we see that

$$\operatorname{cov}(b_1, b_2) \le \big(\operatorname{Var}(b_1)\operatorname{Var}(b_2)\big)^{1/2}.$$

Repeating the argument with $-s_2$ replaced by s_2 shows that

$$\operatorname{cov}(b_1, b_2) \ge -\big(\operatorname{Var}(b_1)\operatorname{Var}(b_2)\big)^{1/2},$$

and these two inequalities together imply (S.07).

> **6.** Let the $n \times k$ matrix $\boldsymbol{X}$ be partitioned as $\boldsymbol{X} = [\boldsymbol{X}_1 \;\; \boldsymbol{X}_2]$, where $\boldsymbol{X}_1$ is $n \times k_1$, $\boldsymbol{X}_2$ is $n \times k_2$, $k_1 + k_2 = k$. For any matrix $\boldsymbol{A}$ with n rows, we denote by $\boldsymbol{P}_{\boldsymbol{A}}$ the orthogonal projection matrix with range the linear span of the columns of $\boldsymbol{A}$, that is $\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{A}(\boldsymbol{A}^\top \boldsymbol{A})^{-1}\boldsymbol{A}^\top$. Let $\boldsymbol{P}_1$ denote $\boldsymbol{P}_{\boldsymbol{X}_1}$. Show that $\boldsymbol{P}_{\boldsymbol{X}} - \boldsymbol{P}_{\boldsymbol{X}_1}$ is an orthogonal projection matrix, that is, show that it is symmetric and idempotent.

The key to this is to notice that

$$\boldsymbol{P}_{\boldsymbol{X}} \boldsymbol{P}_1 = \boldsymbol{P}_1 \boldsymbol{P}_{\boldsymbol{X}} = \boldsymbol{P}_1. \tag{S.08}$$

The first equality follows if we see recall that $\boldsymbol{P_X X}_1 = \boldsymbol{X}_1$ and then observe that

$$\boldsymbol{P_X P}_1 = \boldsymbol{P_X X}_1(\boldsymbol{X}_1^\top \boldsymbol{X})^{-1}\boldsymbol{X}_1^\top = \boldsymbol{X}_1(\boldsymbol{X}_1^\top \boldsymbol{X})^{-1}\boldsymbol{X}_1^\top = \boldsymbol{P}_1.$$

The second equality follows on transposing the first.

The symmetry of $\boldsymbol{P_X} - \boldsymbol{P}_1$ is obvious since both projections are symmetric. Idempotency can now be proved using (S.08):

$$(\boldsymbol{P_X} - \boldsymbol{P}_1)(\boldsymbol{P_X} - \boldsymbol{P}_1) = \boldsymbol{P_X} - \boldsymbol{P}_1 - \boldsymbol{P}_1 + \boldsymbol{P}_1 = \boldsymbol{P_X} - \boldsymbol{P}_1.$$

> Let $\boldsymbol{M}_1 = \boldsymbol{M}_{\boldsymbol{X}_1}$ and $\boldsymbol{P}_1 = \boldsymbol{P}_{\boldsymbol{X}_1}$. Show that $\boldsymbol{P_X} - \boldsymbol{P}_{\boldsymbol{X}_1} = \boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}$, where $\boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}$ is the projection on to the span of $\boldsymbol{M}_1\boldsymbol{X}_2$. This can be done most easily by showing that any vector in $\mathcal{S}(\boldsymbol{M}_1\boldsymbol{X}_2)$ is invariant under the action of $\boldsymbol{P_X} - \boldsymbol{P}_{\boldsymbol{X}_1}$, and that any vector orthogonal to this span is annihilated by $\boldsymbol{P_X} - \boldsymbol{P}_{\boldsymbol{X}_1}$.

Consider a vector in $\mathcal{S}(\boldsymbol{M}_1\boldsymbol{X}_2)$, which we may write as $\boldsymbol{M}_1\boldsymbol{X}_2\boldsymbol{\gamma}$ for some $k_2 \times 1$ vector $\boldsymbol{\gamma}$. Then note that

$$\boldsymbol{P_X M}_1 = \boldsymbol{P_X}(\mathbf{I} - \boldsymbol{P}_1) = \boldsymbol{P_X} - \boldsymbol{P}_1. \tag{S.09}$$

It follows that

$$(\boldsymbol{P_X} - \boldsymbol{P}_1)\boldsymbol{M}_1\boldsymbol{X}_2\boldsymbol{\gamma} = (\boldsymbol{P_X} - \boldsymbol{P}_1)\boldsymbol{X}_2\boldsymbol{\gamma} = \boldsymbol{X}_2\boldsymbol{\gamma} - \boldsymbol{P}_1\boldsymbol{X}_2\boldsymbol{\gamma} = (\mathbf{I} - \boldsymbol{P}_1)\boldsymbol{X}_2\boldsymbol{\gamma} = \boldsymbol{M}_1\boldsymbol{X}_2\boldsymbol{\gamma}.$$

Thus $\boldsymbol{M}_1\boldsymbol{X}_2\boldsymbol{\gamma}$ is invariant under $\boldsymbol{P_X} - \boldsymbol{P}_1$.

Now consider an $n \times 1$ vector $\boldsymbol{z}$ orthogonal to $\mathcal{S}(\boldsymbol{M}_1\boldsymbol{X}_2)$. so that $\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{z} = \mathbf{0}$. Then, using the result (S.09) once more, we see that $(\boldsymbol{P_X} - \boldsymbol{P}_1)\boldsymbol{z} = \boldsymbol{P_X M}_1\boldsymbol{z}$. Write this last expression as

$$\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\begin{bmatrix}\boldsymbol{X}_1^\top \\ \boldsymbol{X}_2^\top\end{bmatrix}\boldsymbol{M}_1\boldsymbol{z}.$$

Since $\boldsymbol{X}_1^\top\boldsymbol{M}_1 = \mathbf{0}$ and since we supposed that $\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{z} = 0$, it follows that the entire expression above is zero, as we wished to show.