Statistics II **Problem Set 2**
Professor: Marcelo J. Moreira **Solutions**
TA: Luan Borelli **August 07, 2024**

---

1. [16.3, LNs] *The matrix* $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ *is full-column rank. Use the fact that* X_1 *is orthogonal to* $X_2^* = M_1 X_2$ *to find*

$$(X'X)^{-1} = \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}^{-1}$$

and sums and products of the terms $(X_1'X_1)^{-1}$, $X_1'X_2$, *and* $(X_2^{*\prime}X_2^*)^{-1}$.

Solution. I shall use the result that for a conformable partitioning $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 \\ \boldsymbol{A}_3 & \boldsymbol{A}_4 \end{bmatrix}$,

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \boldsymbol{A}_1^{-1} + \boldsymbol{A}_1^{-1}\boldsymbol{A}_2(\boldsymbol{A}_4 - \boldsymbol{A}_3\boldsymbol{A}_1^{-1}\boldsymbol{A}_2)^{-1}\boldsymbol{A}_3\boldsymbol{A}_1^{-1} & -\boldsymbol{A}_1^{-1}\boldsymbol{A}_2(\boldsymbol{A}_4 - \boldsymbol{A}_3\boldsymbol{A}_1^{-1}\boldsymbol{A}_2)^{-1} \\ -(\boldsymbol{A}_4 - \boldsymbol{A}_3\boldsymbol{A}_1^{-1}\boldsymbol{A}_2)^{-1}\boldsymbol{A}_3\boldsymbol{A}_1^{-1} & (\boldsymbol{A}_4 - \boldsymbol{A}_3\boldsymbol{A}_1^{-1}\boldsymbol{A}_2)^{-1} \end{bmatrix},$$

provided $\boldsymbol{A}_1$ and the Schur complement of $\boldsymbol{A}_1$ in $\boldsymbol{A}$, $\boldsymbol{A}_4 - \boldsymbol{A}_3\boldsymbol{A}_1^{-1}\boldsymbol{A}_2$, are invertible. Define the Schur complement of $X_1'X_1$ in $(X'X)^{-1}$ as

$$\begin{aligned} S &\equiv X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}X_1'X_2 = X_2'X_2 - X_2'N_1X_2 \\ &= X_2'(I - N_1)X_2 = X_2'M_1X_2 = X_2'M_1M_1X_2 = X_2^{*\prime}X_2^*. \end{aligned}$$

Then we can write

$$\begin{aligned} (X'X)^{-1} &= \begin{bmatrix} (X_1'X_1)^{-1} + (X_1'X_1)^{-1}(X_1'X_2)S^{-1}(X_2'X_1)(X_1'X_1)^{-1} & -(X_1'X_1)^{-1}(X_1'X_2)S^{-1} \\ -S^{-1}X_2'X_1(X_1'X_1)^{-1} & S^{-1} \end{bmatrix} \\ &= \begin{bmatrix} (X_1'X_1)^{-1}[I + (X_1'X_2)S^{-1}(X_1'X_2)'(X_1'X_1)^{-1}] & -(X_1'X_1)^{-1}(X_1'X_2)S^{-1} \\ -[(X_1'X_1)^{-1}(X_1'X_2)S^{-1}]' & S^{-1} \end{bmatrix}, \end{aligned}$$

as desired.

Alternatively, notice that $X_1'M_1 = X_1'(I - X_1(X_1'X_1)^{-1}X_1') = X_1' - X_1'X_1(X_1'X_1)^{-1}X_1' = X_1' - X_1' = 0$. Thus $X_1'X_2^* = X_1'M_1X_2 = 0$. Therefore,

$$\begin{aligned} \left(\begin{bmatrix} X_1 & X_2^* \end{bmatrix}' \begin{bmatrix} X_1 & X_2^* \end{bmatrix}\right)^{-1} &= \begin{bmatrix} X_1'X_1 & X_1'X_2^* \\ X_2^{*\prime}X_1 & X_2^{*\prime}X_2^* \end{bmatrix}^{-1} \\ &= \begin{bmatrix} X_1'X_1 & 0 \\ 0 & X_2^{*\prime}X_2^* \end{bmatrix}^{-1} = \begin{bmatrix} (X_1'X_1)^{-1} & 0 \\ 0 & (X_2^{*\prime}X_2^*)^{-1} \end{bmatrix}. \end{aligned}$$

Now, observe that

$$\begin{aligned} \begin{bmatrix} X_1 & X_2^* \end{bmatrix} \begin{bmatrix} I_{k_1} & (X_1'X_1)^{-1}X_1'X_2 \\ 0 & I_{k_2} \end{bmatrix} &= \begin{bmatrix} X_1 & X_1(X_1'X_1)^{-1}X_1'X_2 + X_2^* \end{bmatrix} \\ &= \begin{bmatrix} X_1 & N_1X_2 + M_1X_2 \end{bmatrix} \\ &= \begin{bmatrix} X_1 & (N_1 + N_1)X_2 \end{bmatrix} \\ &= \begin{bmatrix} X_1 & X_2 \end{bmatrix}, \end{aligned}$$

and let $A \equiv (X_1'X_1)^{-1}X_1'X_2$. Then we can write

$$
\begin{aligned}
(X'X)^{-1} &= \left[\left(\begin{bmatrix} X_1 & X_2^* \end{bmatrix}\begin{bmatrix} I_{k_1} & A \\ 0 & I_{k_2}\end{bmatrix}\right)'\left(\begin{bmatrix} X_1 & X_2^* \end{bmatrix}\begin{bmatrix} I_{k_1} & A \\ 0 & I_{k_2}\end{bmatrix}\right)\right]^{-1} \\
&= \left[\begin{bmatrix} I_{k_1} & A \\ 0 & I_{k_2}\end{bmatrix}'\begin{bmatrix} X_1 & X_2^* \end{bmatrix}'\begin{bmatrix} X_1 & X_2^* \end{bmatrix}\begin{bmatrix} I_{k_1} & A \\ 0 & I_{k_2}\end{bmatrix}\right]^{-1} \\
&= \left[\begin{bmatrix} I_{k_1} & 0' \\ A_{12}' & I_{k_2}\end{bmatrix}\begin{bmatrix} X_1 & X_2^* \end{bmatrix}'\begin{bmatrix} X_1 & X_2^* \end{bmatrix}\begin{bmatrix} I_{k_1} & A \\ 0 & I_{k_2}\end{bmatrix}\right]^{-1} \\
&= \begin{bmatrix} I_{k_1} & A \\ 0 & I_{k_2}\end{bmatrix}^{-1}\left[\begin{bmatrix} X_1 & X_2^* \end{bmatrix}'\begin{bmatrix} X_1 & X_2^* \end{bmatrix}\right]^{-1}\begin{bmatrix} I_{k_1} & 0' \\ A_{12}' & I_{k_2}\end{bmatrix}^{-1} \\
&= \begin{bmatrix} I_{k_1} & -A \\ 0 & I_{k_2}\end{bmatrix}\begin{bmatrix} (X_1'X_1)^{-1} & 0 \\ 0 & (X_2^{*\prime}X_2^*)^{-1}\end{bmatrix}\begin{bmatrix} I_{k_1} & 0' \\ -A_{12}' & I_{k_2}\end{bmatrix} \\
&= \begin{bmatrix} (X_1'X_1)^{-1} & -A(X_2^{*\prime}X_2^*)^{-1} \\ 0 & (X_2^{*\prime}X_2^*)^{-1}\end{bmatrix}\begin{bmatrix} I_{k_1} & 0' \\ -A_{12}' & I_{k_2}\end{bmatrix} \\
&= \begin{bmatrix} (X_1'X_1)^{-1} + A(X_2^{*\prime}X_2^*)^{-1}A' & -A(X_2^{*\prime}X_2^*)^{-1} \\ -(X_2^{*\prime}X_2^*)^{-1}A' & (X_2^{*\prime}X_2^*)^{-1}\end{bmatrix}
\end{aligned}
$$

$$
= \begin{bmatrix} (X_1'X_1)^{-1} + (X_1'X_1)^{-1}X_1'X_2(X_2^{*\prime}X_2^*)^{-1}X_2'X_1(X_1'X_1)^{-1} & -(X_1'X_1)^{-1}X_1'X_2(X_2^{*\prime}X_2^*)^{-1} \\ -(X_2^{*\prime}X_2^*)^{-1}X_2'X_1(X_1'X_1)^{-1} & (X_2^{*\prime}X_2^*)^{-1}\end{bmatrix}.
$$

We obtained the same result, but this time without using the 2×2 block inversion formula.

□

---

2. [16.7, LNs] *Let X and y be*

$$
X = [X_1, X_2] = \begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 3 \\ 1 & 5 \\ 1 & 2 \end{bmatrix} \quad \textit{and} \quad y = \begin{bmatrix} 14 \\ 17 \\ 8 \\ 16 \\ 3 \end{bmatrix}.
$$

Calculate the following:

(a) *$Q = X'X$, $|X'X|$, and Q^{-1}.*

Solution.

$$
Q = \begin{bmatrix} 5 & 16 \\ 16 & 58 \end{bmatrix}, \quad |X'X| = 34, \text{ and } \quad Q^{-1} = \begin{bmatrix} \frac{29}{17} & -\frac{8}{17} \\ -\frac{8}{17} & \frac{5}{34} \end{bmatrix}.
$$

□

(b) *$A = Q^{-1}X'$ and $\hat{\beta} = Ay$.*

Solution.

$$A = \begin{bmatrix} \frac{13}{17} & -\frac{3}{17} & \frac{5}{17} & -\frac{11}{17} & \frac{13}{17} \\ -\frac{3}{17} & \frac{2}{17} & -\frac{1}{34} & \frac{9}{34} & -\frac{3}{17} \end{bmatrix} \quad \text{and } \hat{\beta} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

□

(c) *N and $\hat{y} = Ny$.*

Solution.

$$N = \begin{bmatrix} \frac{7}{17} & \frac{1}{17} & \frac{4}{17} & -\frac{2}{17} & \frac{7}{17} \\ \frac{1}{17} & \frac{5}{17} & \frac{3}{17} & \frac{7}{17} & \frac{1}{17} \\ \frac{4}{17} & \frac{3}{17} & \frac{7}{34} & \frac{5}{34} & \frac{4}{17} \\ -\frac{2}{17} & \frac{7}{17} & \frac{5}{34} & \frac{23}{34} & -\frac{2}{17} \\ \frac{7}{17} & \frac{1}{17} & \frac{4}{17} & -\frac{2}{17} & \frac{7}{17} \end{bmatrix} \quad \text{and} \quad \hat{y} = \begin{bmatrix} 8 \\ 14 \\ 11 \\ 17 \\ 8 \end{bmatrix}.$$

□

(d) *M and $e = My$.*

Solution.

$$M = \begin{bmatrix} \frac{10}{17} & -\frac{1}{17} & -\frac{4}{17} & \frac{2}{17} & -\frac{7}{17} \\ -\frac{1}{17} & \frac{12}{17} & -\frac{3}{17} & -\frac{7}{17} & -\frac{1}{17} \\ -\frac{4}{17} & -\frac{3}{17} & \frac{27}{34} & -\frac{5}{34} & -\frac{4}{17} \\ \frac{2}{17} & -\frac{7}{17} & -\frac{5}{34} & \frac{11}{34} & \frac{2}{17} \\ -\frac{7}{17} & -\frac{1}{17} & -\frac{4}{17} & \frac{2}{17} & \frac{10}{17} \end{bmatrix} \quad \text{and} \quad e = \begin{bmatrix} 6 \\ 3 \\ -3 \\ -1 \\ -5 \end{bmatrix}.$$

□

(e) *$tr(N)$ and $tr(M)$.*

Solution.

$$tr(N) = 2 \quad \text{and} \quad tr(M) = 3.$$

□

(f) *$X_2^{*\prime} X_2^*$, $X_2^{*\prime} X_2$, $X_2^{*\prime} y^*$, and $X_2^{*\prime} y$.*

Solution.

$$X_2^{*\prime} X_2^* = \frac{34}{5}, \quad X_2^{*\prime} X_2 = \frac{34}{5}, \quad X_2^{*\prime} y^* = \frac{102}{5}, \quad \text{and} \quad X_2^{*\prime} y = \frac{102}{5}.$$

□

(g) *$(X_2^{*\prime} X_2^*)^{-1} X_2^{*\prime} y$, and compare your answer with item (b).*

Solution. As in item **(b)** we obtained $\hat{\beta}_2 = 3$, by the FWL theorem the answer must be 3.

$$(X_2^{*\prime} X_2^*)^{-1} X_2^{*\prime} y = 3.$$

□

3. [16.2, LNs] *Let A and B be two symmetric and positive definite $k \times k$ matrices. Prove that $A - B$ is positive definite if and only if $(B^{-1} - A^{-1})$ is positive definite. Apply this result to compare $(X_2^{*\prime} X_2^*)^{-1}$ and $(X_2' X_2)^{-1}$.*

Solution. For this proof we will be extensively using the well-known fact that if $X \succ 0$, then $C' X C \succ 0$ for any conformable C, where "$\succ 0$" here reads "is positive definite". This is just a notation. First, we shall prove that $I_k - A \succ 0 \iff A^{-1} - I_k \succ 0$.

$(\implies)$ Since A is symmetric, so is A^{-1}. Therefore there exists $A^{-1/2}$ symmetric such that $A^{-1} = A^{-1/2} A^{-1/2}$. Notice that $I_k = A^{-1/2} A A^{-1/2}$.[1] We have that

$$A^{-1} - I_k = A^{-1/2} A^{-1/2} - A^{-1/2} A A^{-1/2} = A^{-1/2}(I_k - A) A^{-1/2} \succ 0.$$

The conclusion follows from the fact that $I_k - A \succ 0$, by hypothesis.

$(\impliedby)$ Similarly, since A is symmetric, we can find $A^{1/2}$ such that $A = A^{1/2} A^{1/2}$. Notice that $I_k = A^{1/2} A^{-1} A^{1/2}$. We have that

$$I_k - A = A^{1/2} A^{-1} A^{1/2} - A^{1/2} I_k A^{1/2} = A^{1/2}(A^{-1} - I_k) A^{1/2} \succ 0.$$

The conclusion follows from the fact that $A^{-1} - I_k \succ 0$, by hypothesis.

Now we shall use the above result to prove that $A - B \succ 0 \iff B^{-1} - A^{-1} \succ 0$.

$(\implies)$ Observe that

$$A - B \succ 0 \implies A^{-1/2}(A - B) A^{-1/2} \succ 0 \implies I_k - A^{-1/2} B A^{-1/2} \succ 0.$$

From the previous result, this implies $A^{1/2} B^{-1} A^{1/2} - I_k \succ 0$. Therefore

$$B^{-1} - A^{-1} = A^{-1/2} A^{1/2} B^{-1} A^{1/2} A^{-1/2} - A^{-1/2} I_k A^{-1/2} = A^{-1/2}(A^{1/2} B^{-1} A^{1/2} - I_k) A^{-1/2} \succ 0.$$

$(\impliedby)$ Since B is symmetric, there exists $B^{1/2}$ symmetric such that $B = B^{1/2} B^{1/2}$. We have that $B^{-1} - A^{-1} \succ 0$, whence $B^{1/2}(B^{-1} - A^{-1}) B^{1/2} \succ 0$. Thus $I_k - B^{1/2} A^{-1} B^{1/2} \succ 0$.

The result from the first part implies $B^{-1/2} A B^{-1/2} - I_k \succ 0$. Therefore

$$A - B = B^{1/2} B^{-1/2} A B^{-1/2} B^{1/2} - B^{1/2} I_k B^{1/2} = B^{1/2}(B^{-1/2} A B^{-1/2} - I_k) B^{1/2} \succ 0.$$

Observe that a similar proof to that presented in this exercise could be employed to show that the same result holds for positive *semi*definiteness. Having this in mind, let's compare $(X_2^{*\prime} X_2^*)^{-1}$ and $(X_2' X_2)^{-1}$. Observe that $X_2^{*\prime} X_2^* = X_2' M_1 X_2^* = X_2'(I - N_1) X_2 = X_2' X_2 -$

[1] Observe that

$$A^{-1/2} A A^{-1/2} = A^{-1/2} (A^{-1})^{-1} A^{-1/2} = A^{-1/2} (A^{-1/2} A^{-1/2})^{-1} A^{-1/2} = A^{-1/2} (A^{-1/2})^{-1} (A^{-1/2})^{-1} A^{-1/2} = I_k.$$

$X_2'N_1X_2$. Thus $X_2'X_2 - X_2^{*\prime}X_2^* = X_2'N_1X_2$. Since N_1 is positive semidefinite,[2] it follows that $X_2'N_1X_2$ is positive semidefinite. Therefore $(X_2'X_2)^{-1} - (X_2^{*\prime}X_2^*)^{-1}$ is positive semidefinite. □

---

4. [16.8, LNs] *The regression includes an intercept, and the first column of X is $x_1 = \mathbf{1}$, the vector of ones. Let $M_1 = I - x_1(x_1'x_1)^{-1}x_1'$.*

(a) *Show that M_1y is the vector of residuals from a regression of y on the vector of ones alone.*

Solution. A regression of y on the vector of ones alone gives

$$\hat{\beta} = (\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}'y.$$

Thus the vector of residuals is

$$y - \mathbf{1}\hat{\beta} = y - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}'y = (I - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}')y = (I - x_1(x_1'x_1)^{-1}x_1')y = M_1y.$$

□

(b) *Show that $y'M_1y = \sum_i (y_i - \bar{y})^2$.*

Solution. Observe that

$$\begin{aligned} y'M_1y = y'(I - x_1(x_1'x_1)^{-1}x_1')y &= y'(I - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}')y = y'y - y'\mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}'y = y'y - (\mathbf{1}'y)'(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}'y \\ &= \sum_{i=1}^n y_i^2 - n^{-1}\left(\sum_{i=1}^n y_i\right)^2 = \sum_{i=1}^n y_i^2 - \bar{y}\left(\sum_{i=1}^n y_i\right) = \sum_{i=1}^n (y_i^2 - \bar{y}y_i) = \sum_{i=1}^n (y_i^2 - 2\bar{y}y_i + \bar{y}y_i) \\ &= \sum_{i=1}^n (y_i^2 - 2\bar{y}y_i) + \bar{y}\sum_{i=1}^n y_i = \sum_{i=1}^n (y_i^2 - 2\bar{y}y_i) + n\bar{y}\bar{y} = \sum_{i=1}^n (y_i^2 - 2\bar{y}y_i + \bar{y}^2) \\ &= \sum_i (y_i - \bar{y})^2. \end{aligned}$$

as desired. □

(c) *Suppose the classical regression model applies to $\mathbb{E}[y] = X\beta$. Show that $\mathbb{E}\left[\sum_i (y_i - \bar{y})^2\right] = (n-1)\sigma^2 + \beta_2'X_2^{*\prime}X_2^*\beta_2$, where β_2 is the $(k-1)\times 1$ subvector that remains when the first element of β is deleted, X_2 is the $n \times (k-1)$ submatrix that remains when the first column of X is deleted, and $X_2^* = M_1X_2$.*

[2] Recall that all the eigenvalues of the annihilator matrix are either zero or one; thus, nonnegative. A matrix whose all eigenvalues are nonnegative is always positive semidefinite.

Solution. Observe that

$$\begin{aligned} y'M_1y &= (\mathbf{1}\beta_1 + X_2\beta_2 + u)'M_1(\mathbf{1}\beta_1 + X_2\beta_2 + u) \\ &= (\beta_1'\mathbf{1}' + \beta_2'X_2' + u')M_1M_1(\mathbf{1}\beta_1 + X_2\beta_2 + u) \\ &= (\beta_1'\mathbf{1}'M_1 + \beta_2'X_2'M_1 + u'M_1)(M_1\mathbf{1}\beta_1 + M_1X_2\beta_2 + M_1u) \\ &= (\beta_2'X_2^{*\prime} + u'M_1)(X_2^*\beta_2 + M_1u) \\ &= \beta_2'X_2^{*\prime}X_2^*\beta_2 + \beta_2'X_2^{*\prime}M_1u + u'M_1'X_2^*\beta_2 + u'M_1u. \end{aligned}$$

Thus,

$$\mathbb{E}[y'M_1y] = \mathbb{E}\left[\sum_i (y_i - \bar{y})^2\right] = \beta_2'X_2^{*\prime}X_2^*\beta_2 + \mathbb{E}[u'M_1u] = \beta_2'X_2^{*\prime}X_2^*\beta_2 + \mathbb{E}[\mathrm{tr}(u'M_1u)].$$

Now, recalling that the trace of M_X is $n-k$, where $k = \dim(X)$, we obtain

$$\mathbb{E}[\mathrm{tr}(uu'M_1)] = \mathrm{tr}(\mathbb{E}[uu']M_1) = \mathrm{tr}(I_n\sigma^2M_1) = \mathrm{tr}(M_1)\sigma^2 = (n-1)\sigma^2,$$

whence it follows that $\mathbb{E}[y'M_1y] = \beta_2'X_2^{*\prime}X_2^*\beta_2 + (n-1)\sigma^2$, as desired. □

(d) *Consider $\bar{R}^2 = 1 - [\sum_i e_i^2/(n-k)]\,/\,[\sum_i (y_i-\bar{y})^2/(n-1)]$. Evaluate the claim that the denominator $\sum_i (y_i-\bar{y})^2/(n-1)$ is an unbiased estimator of the variance of the dependent variable.*

Solution. The claim is false. Observe that

$$\mathbb{E}\left[\sum_i (y_i-\bar{y})^2/(n-1)\right] = \frac{\beta_2'X_2^{*\prime}X_2^*\beta_2}{n-1} + \sigma^2 \neq \sigma^2.$$

□

---

5 [16.11, LNs] *Consider the linear regression model*

$$y = X\beta + u,$$

where $\mathbb{E}[u|X] = 0$ and $V(u|x) = \sigma^2 I_n$. There are two regressors, so the i-th row of X is given by $x_t = (x_{1,t}, x_{2,t})$. Let

$$b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \left(\sum_{t=1}^n x_t x_t'\right)^{-1} \sum_{t=1}^n x_t y_t$$

be the OLS coefficient of a regression of y on X.

(a) *Let $b_1^* = \left(\sum_{t=1}^n x_{1,t}^2\right)^{-1} \sum_{t=1}^n x_{1,t}y_t$ be the coefficient of a regression of y on x_1. Show that*

$$b_1^* = b_1 + Fb_2.$$

Solution. Write $y_t = x_1 b_1 + x_{2,t} b_2 + u_t$ and

$$\begin{aligned} b_1^* &= \left(\sum_{t=1}^n x_{1,t}^2\right)^{-1} \sum_{t=1}^n x_{1,t}(x_{1,t}b_1 + x_{2,t}b_2 + u_t) \\ &= b_1 + \left(\sum_{t=1}^n x_{1,t}^2\right)^{-1} \sum_{t=1}^n x_{1,t}x_{2,t}b_2 + \left(\sum_{t=1}^n x_{1,t}^2\right)^{-1} \sum_{t=1}^n x_{1,t}u_t \\ &= b_1 + Fb_2, \end{aligned}$$

where $F \equiv \left(\sum_{t=1}^n x_{1,t}^2\right)^{-1} \sum_{t=1}^n x_{1,t}x_{2,t}$. The last equality follows from the fact that $u_t = My$ and $Mx_{1,t} = 0$. □

(b) *We are given the following estimated OLS coefficients of the dependent variables (columns) and explanatory variables (rows):*

	y	x_1	x_2
y	1		
x_1		1	
x_2			1

Find the estimator $\hat{\beta} = (X'X)^{-1}X'y$.

Solution. It seems that there is information missing in the table. If we assume that the empty entries mean that coefficients are zero, then $x_1'y = x_2'y = x_2'x_1 = 0$. Thus $X'y = (x_1'y, x_2'y)' = (0,0)'$ and hence $\hat{\beta} = (0,0)'$. If, instead, we only assume that the OLS coefficients of x_2 in x_1 or x_1 in x_2 are zero, then we have $x_2'x_1 = x_1'x_2 = 0$. Therefore,

$$(X'X)^{-1} = \left(\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix}\right)^{-1} = \begin{bmatrix} x_1'x_1 & x_1'x_2 \\ x_2'x_1 & x_2'x_2 \end{bmatrix}^{-1} = \begin{bmatrix} x_1'x_1 & 0 \\ 0 & x_2'x_2 \end{bmatrix}^{-1} = \begin{bmatrix} (x_1'x_1)^{-1} & 0 \\ 0 & (x_2'x_2)^{-1} \end{bmatrix},$$

whence it follows that

$$\begin{aligned} \hat{\beta} = (X'X)^{-1}X'y &= \begin{bmatrix} (x_1'x_1)^{-1} & 0 \\ 0 & (x_2'x_2)^{-1} \end{bmatrix} \begin{bmatrix} x_1'y \\ x_2'y \end{bmatrix} \\ &= \begin{bmatrix} (x_1'x_1)^{-1}x_1'y \\ (x_2'x_2)^{-1}x_2'y \end{bmatrix} = \begin{bmatrix} b_1 + F_1 b_2 \\ b_2 + F_2 b_1 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \end{aligned}$$

where, akin to item **(a)**, $F_1 = (x_1'x_1)^{-1} x_1'x_2 = 0$ and $F_2 = (x_2'x_2)^{-1} x_2'x_1 = 0$. □

---

6. *Suppose two researchers are interested in the linear relation between the production of an agricultural product y and fertilizer z. They have quarterly data on these variables from m years and a total of $n = 4m$ observations. The researchers are concerned with seasonal patterns in these variables. Researcher John proposes that first each variable is deseasonalized in the following way: calculate the seasonal means $\bar{y}_1$, $\bar{y}_2$, $\bar{y}_3$, $\bar{y}_4$, and express each observation* *as a deviation from its seasonal mean:* $y^*_{th} = y_{th} - \bar{y}_h$*, where* y_{th} *is the value of* y *for year* t *and quarter* h*, and* $z^*_{th} = z_{th} - \bar{z}_h$*, where* z_{th} *is the value of* z *for year* t *and quarter* h*. Then regress* y^* *on* z^**. On the other hand, researcher Robert propoes to regress* y *on* X *and* z*, where* $X = (x_1, x_2, x_3, x_4)$ *and* x_h *is the* h^{th} *quarter seasonal dummy*

$$x_h = \begin{cases} 1 & \textit{in quarter } h \\ 0 & \textit{otherwise} \end{cases}.$$

(a) *Show that the two competing methods proposed by John and Robert yield the same estimator for the fertilizer effect* z *on* y*.*

Solution. Partition data in quarters: $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4)$ and $\mathbf{z} = (\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4)$, where each $\mathbf{y}_h$ and $\mathbf{z}_h$, $h = 1, ..., 4$ is $m \times 1$. Let $\bar{\mathbf{z}}_h = \mathbf{1}\bar{z}_h$ and $\bar{\mathbf{y}}_h = \mathbf{1}\bar{y}_h$ for $h = 1, \ldots, h$, where $\mathbf{1}$ is a $m \times 1$ vector of ones. John estimates

$$\mathbf{y}^* = \mathbf{z}^*\beta + \mathbf{u},$$

or

$$\begin{bmatrix} \mathbf{y}_1 - \bar{\mathbf{y}}_1 \\ \mathbf{y}_2 - \bar{\mathbf{y}}_2 \\ \mathbf{y}_3 - \bar{\mathbf{y}}_3 \\ \mathbf{y}_4 - \bar{\mathbf{y}}_4 \end{bmatrix} = \begin{bmatrix} \mathbf{z}_1 - \bar{\mathbf{z}}_1 \\ \mathbf{z}_2 - \bar{\mathbf{z}}_2 \\ \mathbf{z}_3 - \bar{\mathbf{z}}_3 \\ \mathbf{z}_4 - \bar{\mathbf{z}}_4 \end{bmatrix} \beta + \mathbf{u}.$$

The OLS estimator for β is given by

$$\hat{\beta} = (\mathbf{z}^{*\prime}\mathbf{z}^*)^{-1}\mathbf{z}^{*\prime}\mathbf{y}^*.$$

Robert estimates

$$\mathbf{y} = \mathbf{z}\beta_1 + X\beta_2 + \mathbf{u}.$$

By FWL, $\hat{\beta}_1 = (\mathbf{z}' M_X \mathbf{z})^{-1}\mathbf{z}' M_X y$, where $M_X = I - X(X'X)^{-1}X'$. Partition

$$X = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix},$$

where $\mathbf{1}$ and $\mathbf{0}$ are $m \times 1$ vectors of ones and zeros, respectively. Observe that

$$\begin{aligned} X'X &= \begin{bmatrix} \mathbf{1}' & \mathbf{0}' & \mathbf{0}' & \mathbf{0}' \\ \mathbf{0}' & \mathbf{1}' & \mathbf{0}' & \mathbf{0}' \\ \mathbf{0}' & \mathbf{0}' & \mathbf{1}' & \mathbf{0}' \\ \mathbf{0}' & \mathbf{0}' & \mathbf{0}' & \mathbf{1}' \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{0}'\mathbf{0} & \mathbf{0}'\mathbf{0} & \mathbf{0}'\mathbf{0} \\ \mathbf{0}'\mathbf{0} & \mathbf{1}'\mathbf{1} & \mathbf{0}'\mathbf{0} & \mathbf{0}'\mathbf{0} \\ \mathbf{0}'\mathbf{0} & \mathbf{0}'\mathbf{0} & \mathbf{1}'\mathbf{1} & \mathbf{0}'\mathbf{0} \\ \mathbf{0}'\mathbf{0} & \mathbf{0}'\mathbf{0} & \mathbf{0}'\mathbf{0} & \mathbf{1}'\mathbf{1} \end{bmatrix} = \begin{bmatrix} m & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & m & 0 \\ 0 & 0 & 0 & m \end{bmatrix} = I_4 m. \end{aligned}$$

Furthermore,

$$XX' = \begin{bmatrix} \mathbf{1}\mathbf{1}' & \mathbf{0}\mathbf{0}' & \mathbf{0}\mathbf{0}' & \mathbf{0}\mathbf{0}' \\ \mathbf{0}\mathbf{0}' & \mathbf{1}\mathbf{1}' & \mathbf{0}\mathbf{0}' & \mathbf{0}\mathbf{0}' \\ \mathbf{0}\mathbf{0}' & \mathbf{0}\mathbf{0}' & \mathbf{1}\mathbf{1}' & \mathbf{0}\mathbf{0}' \\ \mathbf{0}\mathbf{0}' & \mathbf{0}\mathbf{0}' & \mathbf{0}\mathbf{0}' & \mathbf{1}\mathbf{1}' \end{bmatrix}.$$

Therefore

$$\begin{aligned} M_X\mathbf{z} &= \mathbf{z} - m^{-1}XX'\mathbf{z} \\ &= \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \mathbf{z}_3 \\ \mathbf{z}_4 \end{bmatrix} - m^{-1} \begin{bmatrix} \mathbf{1}\mathbf{1}' & \mathbf{0}\mathbf{0}' & \mathbf{0}\mathbf{0}' & \mathbf{0}\mathbf{0}' \\ \mathbf{0}\mathbf{0}' & \mathbf{1}\mathbf{1}' & \mathbf{0}\mathbf{0}' & \mathbf{0}\mathbf{0}' \\ \mathbf{0}\mathbf{0}' & \mathbf{0}\mathbf{0}' & \mathbf{1}\mathbf{1}' & \mathbf{0}\mathbf{0}' \\ \mathbf{0}\mathbf{0}' & \mathbf{0}\mathbf{0}' & \mathbf{0}\mathbf{0}' & \mathbf{1}\mathbf{1}' \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \mathbf{z}_3 \\ \mathbf{z}_4 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \mathbf{z}_3 \\ \mathbf{z}_4 \end{bmatrix} - \begin{bmatrix} m^{-1}\mathbf{1}\mathbf{1}'\mathbf{z}_1 \\ m^{-1}\mathbf{1}\mathbf{1}'\mathbf{z}_2 \\ m^{-1}\mathbf{1}\mathbf{1}'\mathbf{z}_3 \\ m^{-1}\mathbf{1}\mathbf{1}'\mathbf{z}_4 \end{bmatrix} = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \mathbf{z}_3 \\ \mathbf{z}_4 \end{bmatrix} - \begin{bmatrix} \mathbf{1}m^{-1}\sum_{j=1}^m z_{j,1} \\ \mathbf{1}m^{-1}\sum_{j=1}^m z_{j,2} \\ \mathbf{1}m^{-1}\sum_{j=1}^m z_{j,3} \\ \mathbf{1}m^{-1}\sum_{j=1}^m z_{j,4} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \mathbf{z}_3 \\ \mathbf{z}_4 \end{bmatrix} - \begin{bmatrix} \mathbf{1}\bar{z}_1 \\ \mathbf{1}\bar{z}_2 \\ \mathbf{1}\bar{z}_3 \\ \mathbf{1}\bar{z}_4 \end{bmatrix} = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \mathbf{z}_3 \\ \mathbf{z}_4 \end{bmatrix} - \begin{bmatrix} \bar{\mathbf{z}}_1 \\ \bar{\mathbf{z}}_2 \\ \bar{\mathbf{z}}_3 \\ \bar{\mathbf{z}}_4 \end{bmatrix} = \mathbf{z}^*, \end{aligned}$$

where by $z_{j,h}$ we denote the year-j quarter-h observation. Analogously, $M_X\mathbf{y} = \mathbf{y}^*$. Therefore

$$\begin{aligned} \hat{\beta}_1 &= (\mathbf{z}'M_X\mathbf{z})^{-1}\mathbf{z}'M_X\mathbf{y} \\ &= (\mathbf{z}'M_XM_x\mathbf{z})^{-1}\mathbf{z}'M_XM_X\mathbf{y} \\ &= [(M_X\mathbf{z})'M_x\mathbf{z}]^{-1}(M_X\mathbf{z})'M_X\mathbf{y} \\ &= (\mathbf{z}^{*\prime}\mathbf{z}^*)^{-1}\mathbf{z}^{*\prime}\mathbf{y}^* = \hat{\beta}, \end{aligned}$$

whence it follows that the two competing methods proposed by John and Robert yield the same estimator for the effect of z on y. □

(b) *Mark instead suggests the intercept could be relevant, therefore they should regress on* $\mathbf{1} = (1, \ldots, 1)'$, *$X$ and Z. The researchers comment this would not be a good idea. Why?*

Solution. Notice that $\mathbf{1}_{4m} = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 + \mathbf{x}_4$, so if we were to add an intercept we would have perfect multicollinearity. This would imply that the full design matrix $D \equiv \begin{bmatrix} \mathbf{1}_{4m} & \mathbf{z} & \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \mathbf{x}_4 \end{bmatrix}$ would not be full rank and hence the moment matrix $D'D$ would not be invertible. Therefore, the OLS estimator $(D'D)^{-1}D'\mathbf{y}$ would not be well-defined. □

(c) *Now Mark and Robert are interested in testing the hypothesis that there is no seasonal pattern in the data. Mark proposes to regress y on X only and test if the coefficient of X equals to zero. Robert, however, proposes to regress y on X and z and test if the coefficient of X equals to zero. Which of the two methods would you choose? Explain your answer.*

Solution. As $\mathbf{z}$ is relevant for explaining $\mathbf{y}$, omitting it from the regression would imply omitted variable bias. So, in principle, I would prefer Robert's suggestion. We must, however, always be aware of the bias-variance tradeoff when deciding whether or not to include new regressors in a model. □

(d) *The research assistant loads the database on Stata. The database contains the variables y, x_1, x_2, x_3, x_4, z described above. Write the commands he would use to run the regression of y on x_1, x_2, x_3, x_4, z.*

Solution. `regress y x1 x2 x3 x4 z, noconstant` □

---

7. [16.30, LNs] *Consider the model*

$$y_i = x_i'\beta + u_i$$

where (x_i', u_i) are iid with $u_i|x_i$ have the density $f(u) \in C^2$ (with support $-\infty < u < \infty$). Assume that

$$\mathbb{E}[U] = \int_{-\infty}^{\infty} uf(u) = 0$$
$$\text{and } V[U] = \mathbb{E}[U^2] = \int_{-\infty}^{\infty} u^2 f(u) = \sigma^2.$$

(a) *Use transformation of variables to show that the (conditional) pdf of $y_i|x_i$ is given by $g(y_i|x_i) = f(y_i - x_i'\beta)$.*

Solution. Recall that if a continuous random variable X has pdf f_X, then an increasing 1-to-1 transformation $Y = h(X)$ of this random variable has pdf $f_X(h^{-1}(y)) \cdot |\frac{\partial h^{-1}(y)}{\partial y}|$. Here y_i is an increasing one-to-one transformation of u_i, which has density f. The inverse transformation is $h^{-1}(u) = u - x_i'\beta$. Therefore the pdf of y_i is $f(h^{-1}(y_i)) \cdot |\frac{\partial h^{-1}(y_i)}{\partial y_i}| = f(y_i - x_i'\beta) = f(u_i)$. □

(b) *Find the likelihood of $y = (y_1, \ldots, y_n)$ conditional on $X = (x_1, \ldots, x_n)'$.*

Solution. The likelihood is $\mathcal{L}(\beta) = \prod_{i=1}^n f(y_i - x_i'\beta)$. □

(c) *State the Gauss-Markov theorem.*

Solution. In the *homoskedastic* linear regression model, if $\tilde{\beta}$ is a linear unbiased estimator of β, then $V[\tilde{\beta}|X] \geq \sigma^2(X'X)^{-1}$. □

(d) *We will show in item **(f)** that the asymptotic variance of $\sqrt{n}(\tilde{\beta} - \beta^*)$ can be smaller than the asymptotic variance of $\sqrt{n}(\hat{\beta} - \beta^*)$, where $\tilde{\beta}$ is the MLE and $\hat{\beta}$ is the OLS estimator. Explain why this result does not contradict the Gauss-Markov theorem.*

Solution. When the MLE estimator lacks linearity and/or unbiasedness, which is perfectly possible, it falls outside the scope of the Gauss-Markov theorem. Consequently, the asymptotic variance of $\sqrt{n}(\tilde{\beta} - \beta^*)$ can be smaller than that of $\sqrt{n}(\hat{\beta} - \beta^*)$ without posing any contradictions. □

(e) *Find the asymptotic variance of* $\sqrt{n}(\hat{\beta}-\beta^*)$.

Solution. Write $\sqrt{n}(\hat{\beta}-\beta^*) = \left(n^{-1}\sum_{i=1}^n x_i x_i'\right)^{-1}\sqrt{n}\left(n^{-1}\sum_{i=1}^n x_i u_i\right)$. By standard LLN, CMT, CLT, and Slutsky arguments it follows that $\sqrt{n}(\hat{\beta}-\beta^*) \xrightarrow{d} N(0, \mathbb{E}[x_i x_i']^{-1}\sigma^2)$. □

(f) *Show algebraically that (i) the asymptotic variance of* $\sqrt{n}(\tilde{\beta}-\beta^*)$ *is no larger than the asymptotic variance of* $\sqrt{n}(\hat{\beta}-\beta^*)$*; and (ii) give a necessary and sufficient condition on the density* $f(u)$ *for the asymptotic variance of* $\hat{\beta}$ *and* $\tilde{\beta}$ *to be the same.*

Solution. Under standard regularity conditions, taking logs of the likelihood function obtained in **(a)**, using first-order conditions and appealing to the mean value theorem, one can show that $\sqrt{n}(\tilde{\beta}-\beta_0) \xrightarrow{d} N(0, J^{-1})$, where $J = \mathbb{E}\left[\left(\frac{f'(u_i)}{f(u_i)}\right)^2 x_i x_i'\right]$. Recall that $A-B$ is PSD if and only if $B^{-1}-A^{-1}$ is PSD. Therefore

$$\left(\frac{1}{\sigma^2}\mathbb{E}\left[x_i x_i'\right]\right)^{-1} - \left(\mathbb{E}\left[\left(\frac{f'(u_i)}{f(u_i)}\right)^2 x_i x_i'\right]\right)^{-1} \gtrsim 0$$

$$\iff \mathbb{E}\left[\left(\frac{f'(u_i)}{f(u_i)}\right)^2 x_i x_i'\right] - \frac{1}{\sigma^2}\mathbb{E}\left[x_i x_i'\right] \gtrsim 0$$

$$\text{(LIE)} \iff \mathbb{E}\left[\mathbb{E}\left[\left(\frac{f'(u_i)}{f(u_i)}\right)^2 \middle| x_i\right] x_i x_i'\right] - \mathbb{E}\left[\frac{1}{\sigma^2} x_i x_i'\right] \gtrsim 0. \tag{1}$$

From Cauchy-Schwartz inequality,

$$\underbrace{\mathbb{E}[u^2|x_i]}_{=\sigma^2}\mathbb{E}\left[\left(\frac{f'(u_i)}{f(u_i)}\right)^2 \middle| x_i\right] \geq \left(\underbrace{\mathbb{E}\left[u\frac{f'(u_i)}{f(u_i)}\middle| x_i\right]}_{=-1}\right)^2 = 1,^{3}$$

whence

$$\mathbb{E}\left[\left(\frac{f'(u_i)}{f(u_i)}\right)^2 \middle| x_i\right] \geq \frac{1}{\sigma^2}.$$

Therefore (1) holds and hence

$$\text{Avar}(\hat{\beta}) - \text{Avar}(\tilde{\beta}) \gtrsim 0.$$

A necessary and sufficient condition on the density $f(u_i)$ for the asymptotic variance of $\hat{\beta}$ and $\tilde{\beta}$ to be the same is $\mathbb{E}\left[\left(\frac{f'(u_i)}{f(u_i)}\right)^2 \middle| x_i\right] = 1/\sigma^2$. A simple sufficient condition is u_i being normally distributed. Observe that in this case we would have $f'(u_i)/f(u_i) = -u_i/\sigma^2$, whence $\mathbb{E}\left[\left(\frac{f'(u_i)}{f(u_i)}\right)^2 \middle| x_i\right] = 1/\sigma^2$. □

[3] Observe that $\mathbb{E}\left[u_i\frac{f'(u_i)}{f(u_i)}\middle| x_i\right] = \int_{-\infty}^{\infty} u_i \frac{f'(u_i)}{f(u_i)} f(u_i)\, du_i = \int_{-\infty}^{\infty} u_i f'(u_i) du_i = u_i f(u_i)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(u_i)\, du_i$, by integration by parts. Since f is a pdf, the second term equals 1. You can show that the first term is zero.

---

8. [16.11, LNs] Repeated exercise. See Exercise 5.

---

9. [16.26, LNs] *Suppose that the classical normal regression model applies to*

$$E(y) = x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + x_4\beta_4 + x_5\beta_5.$$

A researcher regresses y *on* $(x_1, x_2, x_3, x_4, x_5)$*, and also regresses* w *on* (z_1, z_2)*, where* $w = y - x_1$*,* $z_1 = x_2 - x_4$*, and* $z_2 = x_3$*.*

(a) *State the joint null hypothesis that is testable by comparison of the sum of squared residuals from those two regressions.*

Solution. Write the regression model of y on $(x_1, x_2, x_3, x_4, x_5)$ as

$$y - x_1 = (\beta_1 - 1)x_1 + x_2\beta_2 + x_3\beta_3 + x_4\beta_4 + x_5\beta_5 + u.$$

Observe that $y - x_1 = w$. Further, observe that when $\beta_1 = 1$, $\beta_2 = -\beta_4$, and $\beta_5 = 0$ we have

$$w = (x_2 - x_4)\beta_2 + x_3\beta_3 + u = z_1\beta_2 + z_2\beta_3 + u,$$

which is precisely a regression model for a regression of w on (z_1, z_2). Therefore, the joint null hypothesis that is testable by comparison of the sum of squared residuals from those two regressions is $H_0 : (\beta_1 = 1) \wedge (\beta_2 = -\beta_4) \wedge (\beta_5 = 0)$ against $H_1 : (\beta_1 \neq 1) \vee (\beta_2 \neq -\beta_4) \vee (\beta_5 \neq 0)$, where $\wedge$ and $\vee$ denote the logical "and" and "or" operators, respectively. □

(a) *What is the "numerator degrees of freedom" parameter for that test?*

Solution. The F statistic for the above test is given by

$$F = \frac{(SSR_R - SSR_{UR})/q}{SSR_{UR}/(n-k)},$$

where q is the number of restrictions being tested, n is the sample size, and k the number of regressors. Since errors are normally distributed, F follows an exact $F_{q,n-k}$ distribution, with q "numerator degrees of freedom" and $n - k$ "denominator degrees of freedom". Since the number of restrictions being tested is 3, we conclude that the "numerator degrees of freedom" is 3. □

---

10. [16.27, LNs] *Suppose that the classical normal regression model applies to* $\mathbb{E}[y] = x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + x_4\beta_4$*. Let* $w = y - x_4$*,* $z_1 = x_1$*,* $z_2 = x_2 - x_4$*,* $z_3 = x_3 - x_4$*. For a sample of 104 firms, regression* y *on* (x_1, x_2, x_3, x_4) *gives 70 as the sum of squared residuals, while regression* w *on* (z_1, z_2, z_3) *gives 80 as the sum of squared residuals.*

(a) *Test at the 5% significance level the null hypothesis* $\beta_1 + \beta_2 + \beta_4 = 1$ *against the two-sided alternative* $\beta_2 + \beta_3 + \beta_4 \neq 1$*.*

Solution. We can write the model

$$y = x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + x_4\beta_4 + u$$

as $y - x_4 = x_1\beta_1 + (x_2 - x_4)\beta_2 + (x_3 - x_4)\beta_3 + (\beta_2 + \beta_3 + \beta_4 - 1)x_4 + u$, or, equivalently,

$$w = z_1\beta_1 + z_2\beta_2 + z_3\beta_3 + \gamma x_4 + u,$$

where $\gamma \equiv \beta_2 + \beta_3 + \beta_4 - 1$. Under this equivalent formulation, testing $\beta_1 + \beta_2 + \beta_4 = 1$ against $\beta_1 + \beta_2 + \beta_4 \neq 1$ boils down to testing $\gamma = 0$ against $\gamma \neq 0$. The F-statistic for such a test can be written as

$$F = \frac{SSE(\tilde{\beta}_{CLS}) - SSE(\hat{\beta}_{OLS})}{s^2},$$

where $SSE(\beta) = \sum_{i=1}^{n}(y_i - X_i'\beta)^2$ denote the sum-of-squared errors, $\tilde{\beta}_{CLS}$ is the OLS coefficient obtained from regressing w on (z_1, z_2, z_3) only (i.e., the OLS coefficient obtained under the null restriction $\gamma = 0$), and $\hat{\beta}_{OLS}$ is the OLS coefficient obtained from regressing y on (x_1, x_2, x_3, x_4). At the 5% significance level, we reject the null hypothesis if $F > 3.84$, where 3.84 is the (approximate) 95% quantile of a $\chi^2(1)$ distribution.[4]

We have $SSE(\tilde{\beta}_{CLS}) = 80$, $SSE(\hat{\beta}_{OLS}) = 70$, and $s^2 = 0.7$.[5] It follows that

$$F = \frac{80 - 70}{0.7} \approx 14.29 > 3.84.$$

Therefore, at the 5% significance level we reject the null of $\gamma = \beta_1 + \beta_2 + \beta_4 - 1 = 0$. □

(b) *Let $v = y - x_2$, $t_1 = x_1$, $t_2 = x_3 - x_2$, $t_3 = x_4 - x_2$. If v is regressed on (t_1, t_2, t_3), what sum of squared residuals will be obtained?*

Solution. Observe that $t_1 = z_1$, $t_2 = x_3 - x_4 + x_4 - x_2 = z_3 - z_2$, and $t_3 = -z_2$. Therefore

$$\begin{aligned}
& v = t_1\beta_1 + t_2\beta_2 + t_3\beta_2 + u \\
\iff & y - x_2 = z_1\beta_1 + (z_3 - z_2)\beta_2 - z_2\beta_3 + u \\
\iff & y - x_4 = z_1\beta_1 + (z_3 - z_2)\beta_2 - z_2\beta_3 + x_2 - x_4 + u \\
\iff & w = z_1\beta_1 + z_3\beta_2 - z_2\beta_2 - z_2\beta_3 + z_2 + u \\
\iff & w = z_1\beta_1 + z_2(1 - \beta_2 - \beta_3) + z_3\beta_2 + u \\
\iff & w = z_1\gamma_1 + z_2\gamma_2 + z_3\gamma_3 + u,
\end{aligned}$$

where $\gamma_1 \equiv \beta_1$, $\gamma_2 \equiv 1 - \beta_2 - \beta_3$, and $\gamma_3 \equiv \beta_2$. This is a regression model for a regression of w on (z_1, z_2, z_3). Therefore the sum of squared residuals will be 80. □

[4] Since, for this exercise, the errors are assumed to be normally distributed, it can be shown that the F statistic follows an exact $F_{q,n-k} = F_{1,100}$ distribution. Therefore, we could use the 95% quantile of this distribution, which yields a critical value of approximately 3.94. This value is more precise than the approximate critical value of 3.84 obtained from the asymptotic $\chi^2(1)$ distribution of F. However, given that $F = 14.29$ is a relatively large value, the conclusion would remain the same regardless of the distribution considered.

[5] Recall that

$$s^2 = \frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-k} = \frac{SSE(\hat{\beta}_{OLS})}{n-k}.$$

11. [16.28, LNs] *Consider the following regression model:*

$$y_i = X_i\beta_i + u_i, \quad i = 1, 2$$

where $(y_i)_{n_i\times 1}$*,* $(X_i)_{n_i\times k}$ *are nonrandom,* $(u_i)_{n_i\times 1}$*. Assume* $u_i \sim N(0;\sigma^2 I_{n_i})$ *and that* $E(u_i u_j') = 0$*,* $i \neq j$*, where* i *indicates two groups of a population: married and single individuals. The goal is to test the equality between married's parameters and single's parameters,* $H_0 : \beta_1 = \beta_2$*. Note that you can pile the two equations in just one model:*

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ 0 \end{bmatrix} \beta_1 + \begin{bmatrix} 0 \\ X_2 \end{bmatrix} \beta_2 + u.$$

That is, for $y = (y_1', y_2')'$ *and* $u = (u_1', u_2')'$*, we have*

$$y = \begin{bmatrix} X_1 \\ 0 \end{bmatrix} \beta_1 + \begin{bmatrix} 0 \\ X_2 \end{bmatrix} \beta_2 + u.$$

(a) *Rewrite this model as a function of a new parameter* γ *so that testing* $H_0 : \beta_1 = \beta_2$ *is equivalent to testing* $H_0 : \gamma = 0$*. Compute the test based on the* F*-statistic as the difference between restricted and unrestricted sum of squared residuals:* $SSR_R - SSR_{UR}$*.*

Solution. Write

$$\begin{aligned} y &= \begin{bmatrix} X_1 \\ 0 \end{bmatrix} \beta_1 + \begin{bmatrix} 0 \\ X_2 \end{bmatrix} \beta_2 + u \\ &= \begin{bmatrix} X_1 \\ 0 \end{bmatrix} \beta_1 - \begin{bmatrix} X_1 \\ 0 \end{bmatrix} \beta_2 + \begin{bmatrix} X_1 \\ 0 \end{bmatrix} \beta_2 + \begin{bmatrix} 0 \\ X_2 \end{bmatrix} \beta_2 + u \\ &= \begin{bmatrix} X_1 \\ 0 \end{bmatrix} (\beta_1 - \beta_2) + \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta_2 + u \\ &= \begin{bmatrix} X_1 \\ 0 \end{bmatrix} \gamma + \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta_2 + u, \end{aligned}$$

where $\gamma \equiv \beta_1 - \beta_2$, and $\beta \equiv (\gamma', \beta_2')'$. Testing $H_0 : \beta_1 = \beta_2$ then becomes equivalent to testing $H_0 : \gamma = 0$. The F-statistic for this test is the usual

$$F = \frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-k} = \frac{SSR_R - SSR_{UR}}{SSR_{UR}/(n-k)} = \frac{SSR_R - SSR_{UR}}{s^2},$$

where SSR_{UR} is the sum of squared residuals under the OLS coefficient obtained by regressing y on $\begin{bmatrix} X_1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ X_2 \end{bmatrix}$, and SSR_R is the sum of squared residuals under the OLS coefficient obtained by regressing y on $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$; that is, the constrained least squares estimate under the null restriction $\gamma = 0$. At the 5% significance level, we reject the null hypothesis if $F > 3.84$, where 3.84 is the (approximate) 95% quantile of a $\chi^2(1)$ distribution.[6] □

[6] Again, since errors are assumed to be normal, we could alternatively use the 95% quantile of a $F_{q,n-k}$ distribution, instead of the asymptotic approximate 95% from a $\chi^2(1)$ distribution.

(b) *Show that $SSR_{UR} = SSR_1 + SSR_2$, and that $SSR_R = SSR_3$, where:*

1. *SSR_1 is the SSR obtained from regressing y_1 on X_1.*
2. *SSR_2 is the SSR obtained from regressing y_2 on X_2.*
3. *SSR_3 is the SSR obtained from regressing $(y_1', y_2')'$ on $(X_1', X_2')'$.*

Solution.

$$
\begin{aligned}
SSR_{UR} &= \left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} X_1 \\ 0 \end{bmatrix}\beta_1 - \begin{bmatrix} 0 \\ X_2 \end{bmatrix}\beta_2\right)'\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} X_1 \\ 0 \end{bmatrix}\beta_1 - \begin{bmatrix} 0 \\ X_2 \end{bmatrix}\beta_2\right) \\
&= \left(\begin{bmatrix} y_1' & y_2' \end{bmatrix} - \beta_1'\begin{bmatrix} X_1' & 0' \end{bmatrix} - \beta_2'\begin{bmatrix} 0' & X_2' \end{bmatrix}\right)\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} X_1 \\ 0 \end{bmatrix}\beta_1 - \begin{bmatrix} 0 \\ X_2 \end{bmatrix}\beta_2\right) \\
&= \left(-\beta_1'\begin{bmatrix} X_1' & 0' \end{bmatrix} - \beta_2'\begin{bmatrix} 0' & X_2' \end{bmatrix}\right) \\
&\qquad \times \Bigg(\begin{bmatrix} y_1' & y_2' \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} y_1' & y_2' \end{bmatrix}\begin{bmatrix} X_1 \\ 0 \end{bmatrix}\beta_1 - \begin{bmatrix} y_1' & y_2' \end{bmatrix}\begin{bmatrix} 0 \\ X_2 \end{bmatrix}\beta_2 \\
&\qquad\qquad + \beta_1'\begin{bmatrix} X_1' & 0 \end{bmatrix}'\begin{bmatrix} X_1 \\ 0 \end{bmatrix}\beta_1 + \beta_2'\begin{bmatrix} X_2' & 0 \end{bmatrix}'\begin{bmatrix} X_2 \\ 0 \end{bmatrix}\beta_2\Bigg) \\
&= y_1'y_1 - y_1'X_1\beta_1 + \beta_1'X_1'X_1\beta_1 + y_2'y_2 - y_2'X_2\beta_2 + \beta_2'X_2'X_2\beta_2 \\
&= (y_1 - X_1\beta_1)'(y_1 - X_1\beta_1) + (y_2 - X_2\beta_2)'(y_2 - X_2\beta_2) = SSR_1 + SSR_2.
\end{aligned}
$$

$$
\begin{aligned}
SSR_R &= \left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} X_1 \\ 0 \end{bmatrix}\beta_1 - \begin{bmatrix} 0 \\ X_2 \end{bmatrix}\beta_2\right)'\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} X_1 \\ 0 \end{bmatrix}\beta_1 - \begin{bmatrix} 0 \\ X_2 \end{bmatrix}\beta_2\right) \\
&= \left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} X_1 \\ 0 \end{bmatrix}\beta_2 - \begin{bmatrix} 0 \\ X_2 \end{bmatrix}\beta_2\right)'\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} X_1 \\ 0 \end{bmatrix}\beta_2 - \begin{bmatrix} 0 \\ X_2 \end{bmatrix}\beta_2\right) \\
&= \left(y - \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\beta_2\right)'\left(y - \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\beta_2\right) = SSR_3.
\end{aligned}
$$

□