Econ 7010 - Final 2019 Answers

Ryan T. Godwin

1. We should use the GLS estimator when the relative variances of the observations are *known* (when the ω_i are known). There is one common case where the ω_i are indeed known: when the observations are averaged over groups and the size of the groups are known. The advantage is that the estimated standard errors are consistent without having to use any "robust" estimator, and GLS is efficient relative to OLS.

2. The RLS estimator is (i) unbiased when the restrictions are true; (ii) efficient compared to OLS no matter what.

3. Let the model be:

$$\boldsymbol{y} = \boldsymbol{i}\boldsymbol{\beta}_1 + X\boldsymbol{\beta}_2 + \boldsymbol{\epsilon}$$

where $\boldsymbol{i}$ is a column of 1s. By the FWL theorem the LS estimators are:

$$\boldsymbol{b}_1 = \left(\boldsymbol{i}'M_X\boldsymbol{i}\right)^{-1}\boldsymbol{i}'M_X\boldsymbol{y}$$
$$\boldsymbol{b}_2 = \left(X'M_iX\right)^{-1}X'M_i\boldsymbol{y}$$

where $M_X = I - X(X'X)^{-1}X'$, and $M_i = I - \boldsymbol{i}(\boldsymbol{i}'\boldsymbol{i})^{-1}\boldsymbol{i}'$. Now,

$$E[\boldsymbol{b}_1] = \boldsymbol{\beta}_1 + \left(\boldsymbol{i}'M_X\boldsymbol{i}\right)^{-1}\boldsymbol{i}'M_X\boldsymbol{c} \neq \boldsymbol{\beta}_1$$

and

$$E[\boldsymbol{b}_2] = \boldsymbol{\beta}_2 + \left(X'M_iX\right)^{-1}X'M_i\boldsymbol{c} = \boldsymbol{\beta}_2$$

since $M_i\boldsymbol{c} = 0$.

4. Removing a regressors, say X_j, is equivalent to setting the constraint that $\boldsymbol{\beta}_j = \boldsymbol{0}$. Minimizing subject to a constraint cannot improve the minimization. Since LS minimizes the sum of squared residuals by choosing the $\boldsymbol{\beta}$, imposing a constraint on $\boldsymbol{\beta}$ will typically increase the minimized value of $\boldsymbol{e}'\boldsymbol{e}$.

5. See this proof in Section 8.2.

6. See Section 12.4.

7. See the first part of Section 8.3.

8. The GLS estimator is $\widehat{\boldsymbol{\beta}}_{GLS} = \left[X'\Omega^{-1}X\right]^{-1}X'\Omega^{-1}\boldsymbol{y}$. When each observation has been averaged over a group j of size n_j ($j = 1 \ldots m$) then the Ω matrix is:

$$\begin{bmatrix} 1/n_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1/n_m \end{bmatrix}$$

9. See the 2020 Final exam, question 1.

10. See the 2020 Final exam, question 4.

11.

12.

13. a) When the model is non-linear in the parameters, the FOCs to solve for the NLS estimator are also non-linear and cannot be solved analytically (there is no "closed-form" solution). Hence, to locate the minimum of the sum of squared residuals, a numerical algorithm must be used.

    b)

    c) (i) The algorithm may fail to locate a global minimum, instead finding a local minimum. (ii) The algorithm may oscillate (cycle), where the next iteration leads to similar parameter values of a previous iteration. (iii) The algorithm may find a point where the Hessian is singular (the slope of the gradient is near zero). In the case, the algorithm fails.

    d)

$$f(\theta) = \theta^3 - 3\theta$$

$$g(\theta) = 3\theta^2 - 3$$

$$H(\theta) = 6\theta$$

$$\theta_1 = 2 - \left(\frac{9}{12}\right) = 1.25$$

$$\theta_2 = 1.25 - \left(\frac{1.6875}{7.5}\right) = 1.025$$

$$\theta_3 = 1.025 - \left(\frac{0.151875}{6.15}\right) = 1.000305$$

The algorithm appears to be approaching 1. (Setting $g(\theta) = 0$ we can see that 1 is indeed a solution).