# ECOM40006/90013 Econometrics 3

## Week 7 Extras

### Question 1: Ordinary Least Squares

We have already done a fair amount of work on the OLS framework in its stacked data representation. Now, let's take the time to revisit it from an individual data vector perspective. This will be expanded on later in the next question. Suppose that you are dealing with a linear model of the form

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

where $x_i$ and $\beta$ are $k \times 1$ column vectors, and $y_i$ and $u_i$ are both scalars.

**(a)** Derive the score for the mimimization problem

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{n} u_i^2 = \arg\min_{\beta} \sum_{i=1}^{n} (y_i - x_i'\beta)^2$$

The Chain Rule might help for obtaining the score, along with the vector calculus rule

$$\frac{\partial x_i'\beta}{\partial \beta} = x_i.$$

**(b)** Carefully write out the exact expression for the first-order condition (FOC) with respect to (w.r.t.) $\beta$. Show that the score can be written in the form

$$s_n(\hat{\beta}) = \sum_{i=1}^{n} \left.\frac{\partial x_i'\beta}{\partial \beta}\right|_{\beta=\hat{\beta}} (y_i - x_i\hat{\beta}) = \sum_{i=1}^{n} \text{deriv. of mean} \times \text{residuals}$$

**(c)** Solve the first-order condition in (b) for the OLS estimator $\hat{\beta}$.

### Question 2: Weighted Least Squares

Consider again the same model as in Question 1, except that the assumptions on the disturbance term $u_i$ are now slightly different:

$$y_i = x_i'\beta + u_i, \qquad u_i|x_i \sim N(0, f(x_i))$$

That is: this linear model exhibits heteroskedasticity because the conditional variance of $u_i$ depends on the regressors $x_i$ in some way, which we'll call $f(x_i)$. Having an exact form isn't necessary here – the important part is knowing that the heteroskedasticity takes on the form $f(x_i)$.

**(a)** Explain whether there is any change to the form of the OLS estimator obtained from solving the same problem as in Question 1, i.e.

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{n} u_i^2.$$

(Hint: There is no change... but why?)

**(b)** Suppose that we multiply both sides of the linear model by $1/\sqrt{f(x_i)}$ so that one has

$$\begin{aligned} \frac{y_i}{\sqrt{f(x_i)}} &= \frac{x_i'}{\sqrt{f(x_i)}}\beta + \frac{u_i}{\sqrt{f(x_i)}} \\ \implies y_i^* &= x_i^{*\prime}\beta + u_i^*, \end{aligned}$$

where

$$y_i^* = \frac{y_i}{\sqrt{f(x_i)}}, \qquad x_i^* = \frac{x_i}{\sqrt{f(x_i)}}, \qquad u_i^* = \frac{u_i}{\sqrt{f(x_i)}}.$$

Calculate the conditional variance $\mathrm{Var}(u_i^*|x_i)$. Is heteroskedasticity present in this model? (Hint: no.)

**(c)** Write out the OLS estimator from regressing $y_i^*$ on $x_i^*$ in terms of the untransformed data $y_i$, $x_i$ and the form of heteroskedasticity $f(x_i)$.

- Hint: you know how we solved for the OLS estimator earlier? Try using that formula, but replace $x_i$ and $y_i$ with $x_i^*$ and $y_i^*$, then substitute in their definitions. This question isn't intended to take a very long time – the necessary derivations are something we'll talk about shortly.

**(d)** Write out the first-order conditions associated with obtaining the OLS estimator above. Specifically, begin with the minimization problem

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{n} u_i^{*2}$$

and show that the FOC with respect to $\beta$ can be written to be identical to the first-order conditions from Question 1(b), except that there is a variance weighting. That is:

$$\begin{aligned} s_n(\hat{\beta}) &= \sum_{i=1}^{n} \left.\frac{\partial x_i'\beta}{\partial \beta}\right|_{\beta=\hat{\beta}} \times \frac{y_i - x_i\hat{\beta}}{f(x_i)} \\ &= \sum_{i=1}^{n} \text{deriv. of mean} \times \frac{\text{residuals}}{\text{variance weighting}} = 0 \end{aligned}$$

Can you see where the *weighted*, or *generalized* least squares setup comes from when looking at the first-order conditions?

## Question 3: The Feasible GLS Estimator

Consider a linear model $y = X\beta + u$, where the conditional distribution of $u$ given $X$ is known: namely $u|X \sim N(0, \Omega)$. Consider

$$\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y.$$

**(a)** As revision, describe briefly how the GLS estimator can be derived.

The weighting matrix $\Omega^{-1}$ is generally not observable in data, so the *Feasible GLS* (FGLS) method exists as an alternative to GLS estimation. The idea is that $\Omega$ is replaced by an estimator for $\Omega$ that does not need to be true, then the GLS estimator is computed using this replacement for $\Omega$.

Suppose that an estimator $S = S(\beta)$ is estimated for $\Omega$, with estimated value $\hat{S}$ which is symmetric and invertible. This implies the FGLS estimator

$$\hat{\beta}_{FGLS} = (X'\hat{S}^{-1}X)^{-1}X'\hat{S}^{-1}y.$$

For this question, you will need the assumptions that

$$\frac{1}{N}X'\hat{S}^{-1}X \xrightarrow{p} Q \text{ symmetric}, \qquad \frac{1}{\sqrt{N}}X'\hat{S}^{-1}u \xrightarrow{d} N(0, P).$$

**(b)** Derive the conditional variance of the FGLS estimator, namely $\Omega_{FGLS} = \text{Var}(\hat{\beta}_{FGLS}|X)$.

**(c)** Use the variance decomposition formula

$$\text{Var}(X) = \mathbb{E}(\text{Var}(X|Y)) + \text{Var}(\mathbb{E}(X|Y))$$

to show that the unconditional variance of $X'\hat{S}^{-1}u$ can in fact be written as

$$P = \text{Var}(X'\hat{S}^{-1}u) = \mathbb{E}(X'\hat{S}^{-1}\Omega\hat{S}^{-1}X).$$

*Note: You could also use the formula for unconditional variance, which might be simpler.*

**(d)** Show that the distribution of the FGLS estimator can be written as

$$\sqrt{N}(\hat{\beta}_{FGLS} - \beta) \xrightarrow{d} N(0, Q^{-1}PQ^{-1}).$$

**(e)** If the FGLS variance estimator is consistent i.e. $\hat{S} \xrightarrow{p} \Omega$, what happens to the asymptotic distribution of the FGLS estimator? (*Hint: write $Q$ in terms of expectations, then check what $P$ converges to and compare it to $Q$.*)