Supplemental Materials to Lab 9 of QM2

Shane Xinyang Xuan
ShaneXuan.com

Department of Political Science
University of California, San Diego

March 9, 2018

1 Generalized Least Squares

OLS is efficient if $\text{var}[u|\mathbf{X}] = \sigma^2\mathbf{I}$. In the case in which errors are *not* i.i.d., we may assume $\text{var}[u|\mathbf{X}] = \mathbf{\Omega}$, where $\mathbf{\Omega}$ is a known and nonsingular error variance matrix, $\mathbf{\Omega}^{1/2}\mathbf{\Omega}^{1/2} = \mathbf{\Omega}$.

Consider the linear model

$$y = \mathbf{X}\beta + u$$

We premultiply the equation by $\mathbf{\Omega}^{-1/2}$:

$$\mathbf{\Omega}^{-1/2}y = \mathbf{\Omega}^{-1/2}\mathbf{X}\beta + \mathbf{\Omega}^{-1/2}u \tag{1}$$

We care about

$$\begin{aligned}
\text{var}\left[\mathbf{\Omega}^{-1/2}u|\mathbf{X}\right] &= \text{var}[\mathbf{\Omega}^{-1/2}y - \mathbf{\Omega}^{-1/2}\mathbf{X}\beta|\mathbf{X}] &(2)\\
&= \text{var}[\mathbf{\Omega}^{-1/2}u|\mathbf{X}] &(3)\\
&= \mathbb{E}\left[\underbrace{(\mathbf{\Omega}^{-1/2}u)}_{N\times 1}\underbrace{(\mathbf{\Omega}^{-1/2}u)'}_{1\times N}|\mathbf{X}\right] &(4)\\
&= \mathbf{I}_{(N\times N)} &(5)
\end{aligned}$$

Note that the errors are zero mean, uncorrelated, and homoskedastic; β can be efficiently estimated by OLS of $\mathbf{\Omega}^{-1/2}y$ on $\mathbf{\Omega}^{-1/2}\mathbf{X}$:

$$\begin{aligned}
\hat{\beta}_{\text{GLS}} &= \left(\mathbf{X}'\mathbf{\Omega}^{-1/2}\mathbf{\Omega}^{-1/2}\mathbf{X}\right)^{-1}\left(\mathbf{X}'\mathbf{\Omega}^{-1/2}\right)\left(\mathbf{\Omega}^{-1/2}y\right) &(6)\\
&= \left(\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{\Omega}^{-1}y &(7)
\end{aligned}$$

Since $\mathbf{\Omega}$ is unknown, we can model Eq (7) with $\hat{\mathbf{\Omega}}$:

$$\hat{\beta}_{\text{FGLS}} = \left(\mathbf{X}'\hat{\mathbf{\Omega}}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\hat{\mathbf{\Omega}}^{-1}y \tag{8}$$

There are many ways to model heteroskedasticity, and one way is to assume that

$$\text{var}\left[u|\mathbf{X}\right] = u^2 = \sigma^2\exp(\delta_0 + \delta_1 x_1 + ... + \delta_k x_k) \tag{9}$$

It follows that

$$g \equiv \log\left(\hat{u}^2\right) = \alpha_0 + \delta_1 x_1 + ... + \delta_k x_k + e \tag{10}$$

$$h \equiv \exp(\hat{g}) = \widehat{(\hat{u}^2)} \tag{11}$$

To implement this in `Stata`,

- Regress y on $\mathbf{X}$
- Predict $\hat{u}$ and generate $g := \log\left(\hat{u}^2\right)$
- Regress g on $\mathbf{X}$
- Predict $\hat{g}$ and generate $h := \exp(\hat{g})$
- Estimate using weight of $\frac{1}{\hat{h}}$

2 Appendix: Code

Posted on the course website.