BOSTON COLLEGE
Department of Economics
EC771: Econometrics
Spring 2004
Prof. Baum, Ms. Uysal

Solution Key for Problem Set 6

1. What is the covariance matrix, cov$[\hat{\beta},\, \hat{\beta} - \mathbf{b}]$, of the GLS estimator $\hat{\beta} = (\mathbf{X'\Omega^{-1}X})^{-1}\mathbf{X'\Omega^{-1}y}$ and the difference between it and the OLS estimator, $\mathbf{b} = (\mathbf{X'X})^{-1}\mathbf{X'y}$? The result plays a pivotal role in the development of specification tests in Hausman(1978).

Write the two estimators as $\hat{\beta}$=$\beta + (\mathbf{X'\Omega^{-1}X})^{-1}\mathbf{X'\Omega^{-1}}\epsilon$ and $\mathbf{b}$=$\beta + (\mathbf{X'X})^{-1}\mathbf{X'}\epsilon$. Then, $(\hat{\beta} - \mathbf{b}) = [(\mathbf{X'\Omega^{-1}X})^{-1}\mathbf{X'\Omega^{-1}} - (\mathbf{X'X})^{-1}\mathbf{X'}]\epsilon$ has $E[\hat{\beta} - \mathbf{b}] = \mathbf{0}$ since both estimators are unbiased. Therefore, Cov$[\hat{\beta},\, \hat{\beta} - \mathbf{b}] = E[(\hat{\beta} - \beta)(\hat{\beta} - \mathbf{b})']$. Then,

$$E\{(\mathbf{X'\Omega^{-1}X})^{-1}\mathbf{X'\Omega^{-1}}\epsilon\epsilon'[(\mathbf{X'\Omega^{-1}X})^{-1}\mathbf{X'\Omega^{-1}} - (\mathbf{X'X})^{-1}\mathbf{X'}]'\}$$

$$\begin{aligned}
&= (\mathbf{X'\Omega^{-1}X})^{-1}\mathbf{X'\Omega^{-1}}(\sigma^{\mathbf{2}}\mathbf{\Omega})[\mathbf{\Omega^{-1}X}(\mathbf{X'\Omega^{-1}X})^{-1} - \mathbf{X}(\mathbf{X'X})^{-1}] \\
&= \sigma^{\mathbf{2}}(\mathbf{X'\Omega^{-1}X})^{-1}\mathbf{X'\Omega^{-1}\Omega\Omega^{-1}X}(\mathbf{X'\Omega^{-1}X})^{-1} - (\mathbf{X'\Omega^{-1}X})^{-1}\mathbf{X'\Omega^{-1}\Omega X}(\mathbf{X'X})^{-1} \\
&= (\mathbf{X'\Omega^{-1}X})^{-1}(\mathbf{X'\Omega^{-1}X})(\mathbf{X'\Omega^{-1}X})^{-1} - (\mathbf{X'\Omega^{-1}X})^{-1}(\mathbf{X'X})(\mathbf{X'X})^{-1} = \mathbf{0}
\end{aligned}$$

once the inverse matrices are mutiplied.

2. Suppose that the regression model is $y = \mu + \epsilon$, where ϵ has a zero mean, constant variance, and equal correlation ρ across observations. Then $cov[\epsilon_i, \epsilon_j] = \sigma^2\rho$ if $i \neq j$. Prove that the least squares estimator of μ is inconsistent. Find the characteristic roots of $\mathbf{\Omega}$ and show that Condition 2 after Theorem 10.2 is violated.

The covariance matrix is

$$\sigma^2\mathbf{\Omega} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & 1 & \cdots & \rho \\ & & & \vdots & \\ \rho & \rho & \rho & \cdots & 1 \end{bmatrix}$$

The matrix $\mathbf{X}$ is a column of 1s, so the least squares estimator of μ is $\bar{y}$. Inserting this $\mathbf{\Omega}$ into (10-5), we obtain var$[\bar{y}]$=$\frac{\sigma^2}{n}(1 - \rho + n\rho)$. The limit of this expression is $\rho\sigma^{\mathbf{2}}$, not zero. Although ordinary least squares is unbiased, it is not consistent. For this model, $(\mathbf{X'\Omega X})$/n=1 + $\rho(n-1)$, which does not converge. Using theorem 10.2 instead, $\mathbf{X}$ is a column of 1s, so $(\mathbf{X'X})$=n, a scalar, which satisfies condition 1. To find the characteristic roots, multiply out the equation $\mathbf{\Omega X} = \lambda\mathbf{x} = (\mathbf{1} - \rho)\mathbf{Ix} + \rho\mathbf{ii'x} = \lambda\mathbf{x}$. Since $\mathbf{i'x} = \sum_i x_i$, consider any vector $\mathbf{x}$ whose elements sum to zero. If so, then it's obvious that $\lambda = \rho$. There are $n-1$ such roots. Finally, suppose that $\mathbf{x} = \mathbf{i}$. Plugging this into the equation produces $\lambda = 1 - \rho + n\rho$. The characteristic roots of $\mathbf{\Omega}$ are $(1-\rho)$ with multiplicity $n-1$ and $(1-\rho+n\rho)$, which violates condition 2.

3. Suppose that the regression model is $y_i = \mu + \epsilon_i$, where E$[\epsilon_i|x_i]$=0, but var$[\epsilon_i|x_i]$=$\sigma^2 x_i^2$, $x_i > 0$.
(a) Given a sample of observations on y_i and $x-i$, what is the most efficient estimator of ϵ? What is its variance?
(b) What is the ordinary least squares estimator of μ and what is the variance of the ordinary least squares estimator?
(c) Prove that the estimator in (a) is at least as efficient as the estimator in (b).

This is a heteroskedastic regression model in which the matrix $\mathbf{X}$ is a column of ones. The efficient estimator is the GLS estimator, $\hat{\beta} = (\mathbf{X'\Omega^{-1}X})^{-1}\mathbf{X'\Omega^{-1}y} = [\sum_i 1y_i/x_i^2]/[\sum_i 1^2/x_i^2] = [\sum_i(y_i/x_i^2)]/[\sum_i(1/x_i^2)]$. As always, the variance of the estimator is var$[\hat{\beta}]$=$\sigma^2(\mathbf{X'\Omega^{-1}X})^{-1}$=$\sigma^2/[\sum_i(1/x_i^2)]$. The OLS estimator is $(\mathbf{X'X})^{-1}\mathbf{X'y} = \bar{\mathbf{y}}$. The variance of $\bar{y}$ is $\sigma^2(\mathbf{X'X})^{-1}(\mathbf{X'\Omega X})(\mathbf{X'X})^{-1} = (\sigma^2/n^2)\sum_i x_i^2$. To show that the variance of the OLS estimator is greater than or equal to that of the GLS estimator, we must show that $(\sigma^2/n^2)\sum_i x_i^2 \geq \sigma^2/[\sum_i(1/x_i^2)]$ or $(1/n^2)(\sum_i x_i^2)[\sum_i(1/x_i^2)] \geq 1$ or $\sum_i\sum_j(x_i^2/x_j^2) \geq n^2$. The double sum contains n terms equal to one. There remain $n(n-1)/2$ pairs of the form $(x_i^2/x_j^2 + x_j^2/x_i^2)$. If it can be shown that each of these sums is greater than or equal to 2, the result is proved. Just let $z_i = x_i^2$. Then, we require $z_i/z_j + z_j/z_i - 2 \geq 0$. But this is equivalent to $(z_i^2 + z_j^2 - 2z_iz_j)/(z_iz_j) \geq 0$ or $(z_i - z_j)^2/(z_iz_j) \geq 0$, which is certainly true if z_i and z_j are positive. They are since z_i equals x_i^2. This completes the proof.

5. Does first differencing reduce autocorrelation? Consider the models $y_t = \beta'\mathbf{x_t} + \epsilon_t$, where $\epsilon_t = \rho\epsilon_{t-1} + u_t$ and $\epsilon_t = u_t - \lambda u_{t-1}$. Compare the autocorrelation of ϵ_t in the original model to that of ν_t in $y_t - y_{t-1} = \beta'(\mathbf{x_t} - \mathbf{x_{t-1}}) + \nu_t$ where $\nu_t = \epsilon_t - \epsilon_{t-1}$.

For the first order autoregressive model, the autocorrelation is ρ. Consider the first difference, $\nu_t = \epsilon_t - \epsilon_{t-1}$ which has var$[\nu_t]$=2var$[\epsilon_t]$-2cov$[(\epsilon_t, \epsilon_{t-1})] = 2\sigma_\mu^2[1/(1-\rho^2) - \rho/(1-\rho^2)] = 2\sigma_\mu^2/(1+\rho)$ and cov$[\nu_t, \nu_{t-1}]$ = 2cov$[\epsilon_t, \epsilon_{t-1}]$-var$[\epsilon_t]$-cov$[\epsilon_t, \epsilon_{t-1}]$=$\sigma_\mu^2[1/(1-\rho^2)][2\rho - 1 - \rho^2]$=$\sigma_\mu^2[(\rho-1)/(1+\rho)]$. Therefore, the autocorrelation of the differenced process is cov$[\nu_t, \nu_{t-1}]$/var$[\nu_t]$=$(\rho-1)/2$. First differencing reduces the absolute value of the autocorrelation coefficient when ρ is greater than 1/3. For economic data, this is likely to be fairly common.

For the moving average process, the first order autocorrelation is cov$[\epsilon_t, \epsilon_{t-1}]$/var$[\epsilon_t]$ = $-\lambda/(1+\lambda^2)$. To obtain the autocorrelation of the first difference, write $\epsilon_t - \epsilon_{t-1} = u_t - (1+\lambda)u_{t-1} + \lambda u_{t-2}$ and $\epsilon_{t-1} - \epsilon_{t-2} = u_{t-1} - (1+\lambda)u_{t-2} + \lambda u_{t-3}$. The variance of the difference is var$[\epsilon_t - \epsilon_{t-1}]$=$\sigma_\mu^2[(1+\lambda)^2 + (1+\lambda^2)]$. The covariance can be found by taking the expected product of terms with equal subscripts. Thus, $\mathrm{cov}[\epsilon_t - \epsilon_{t-1}, \epsilon_{t-1} - \epsilon_{t-2}] = -\sigma_\mu^2(1+\lambda)^2$. The autocorrelation is $\mathrm{cov}[\epsilon_t - \epsilon_{t-1}, \epsilon_{t-1} - \epsilon_{t-2}]/\mathrm{var}[\epsilon_t - \epsilon_{t-1}] = -(1+\lambda)^2/[(1+\lambda)^2 + (1+\lambda^2)]$. For most of the range of the autocorrelation of the original series, differences increases autocorrelation. But, for most of the range of values that are economically meaningful, differencing reduces autocorrelation.