Problem Set 4 – Solutions

ECON 837

Brian Krauth (adapted from problems by Simon Woodcock), Spring 2010

1. We have $\mathbf{Z} = [\mathbf{y}\ \mathbf{X}]$, so that

$$\mathbf{Z}'\mathbf{Z} = \begin{bmatrix} \mathbf{y}' \\ \mathbf{X}' \end{bmatrix} \begin{bmatrix} \mathbf{y} & \mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{y}'\mathbf{y} & \mathbf{y}'\mathbf{X} \\ \mathbf{X}'\mathbf{y} & \mathbf{X}'\mathbf{X} \end{bmatrix}$$

and hence

$$\mathbf{y}'\mathbf{y} = 150 \quad \mathbf{X}'\mathbf{y} = \begin{bmatrix} 15 \\ 50 \end{bmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{bmatrix} 25 & 0 \\ 0 & 100 \end{bmatrix}.$$

Since the model includes an intercept, we know the $(1,1)$ element of $\mathbf{X}'\mathbf{X}$ is n, and the first element of $\mathbf{X}'\mathbf{y}$ is $\sum y_i$. Therefore

$$\begin{aligned}
\hat{\beta} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \begin{bmatrix} 1/25 & 0 \\ 0 & 1/100 \end{bmatrix} \begin{bmatrix} 15 \\ 50 \end{bmatrix} = \begin{bmatrix} 3/5 \\ 1/2 \end{bmatrix} \\
s^2 &= \frac{\mathbf{e}'\mathbf{e}}{n-k} = \frac{1}{25-2}\left(\mathbf{y} - \mathbf{X}\hat{\beta}\right)'\left(\mathbf{y} - \mathbf{X}\hat{\beta}\right) = \frac{1}{23}\left(\mathbf{y}'\mathbf{y} - 2\mathbf{y}'\mathbf{X}\hat{\beta} + \hat{\beta}'\mathbf{X}'\mathbf{X}\hat{\beta}\right) \\
&= \frac{1}{23}\left(150 - 2\begin{bmatrix} 15 & 50 \end{bmatrix}\begin{bmatrix} 3/5 \\ 1/2 \end{bmatrix} + \begin{bmatrix} 3/5 & 1/2 \end{bmatrix}\begin{bmatrix} 25 & 0 \\ 0 & 100 \end{bmatrix}\begin{bmatrix} 3/5 \\ 1/2 \end{bmatrix}\right) \\
&= \frac{1}{23}\left(150 - 2(9+25) + (9+25)\right) = 116/23 \\
R^2 &= 1 - \frac{\mathbf{e}'\mathbf{e}}{\mathbf{y}'\mathbf{J}\mathbf{y}} = 1 - \frac{\mathbf{e}'\mathbf{e}}{\mathbf{y}'\mathbf{y} - \frac{1}{n}\mathbf{y}'\mathbf{i}\mathbf{i}'\mathbf{y}} = 1 - \frac{116}{150 - \frac{1}{25}\left(\sum y_i\right)^2} = 1 - \frac{116}{150 - 15^2/25} \\
&= 1 - \frac{116}{141} = 25/141 \approx 0.177
\end{aligned}$$

So the cross-products matrix is sufficient for computing all of these, plus a few other things (test of $H_0 : \beta_2 = 0$, for example). However, there are other things we can/should do that we cannot using only the cross-products, such as plot residuals for evidence of heteroskedasticity, etc.

2. We know that in model (1),

$$\begin{matrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{matrix} = \begin{bmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\mathbf{X}_2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}_1'\mathbf{y} \\ \mathbf{X}_2'\mathbf{y} \end{bmatrix}.$$

Applying the partitioned inverse formula, we find the upper left block of the inverse cross-products matrix is $(\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1)^{-1}$, and the upper right block is $-(\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1}$. Therefore

$$\begin{aligned}
\hat{\beta}_1 &= (\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y} - (\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{y} \\
&= (\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1)^{-1}\mathbf{X}_1'\left(\mathbf{I}_n - \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'\right)\mathbf{y} \\
&= (\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{M}_2\mathbf{y}.
\end{aligned}$$

In model (2), we have

$$\begin{aligned}
\tilde{\beta}_1 &= (\mathbf{X}_1'\mathbf{M}_2'\mathbf{M}_2\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{M}_2'\mathbf{M}_2\mathbf{y} \\
&= (\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{M}_2\mathbf{y} = \hat{\beta}_1.
\end{aligned}$$

3.

   (a) We know that

$$\begin{aligned}
\hat{\beta} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\
\hat{\gamma} &= (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}.
\end{aligned}$$

   Since $\mathbf{Z} = \mathbf{X}\mathbf{G}$ and $\mathbf{G}$ is non-singular, we have

$$\begin{aligned}
\hat{\gamma} &= \left[(\mathbf{X}\mathbf{G})'(\mathbf{X}\mathbf{G})\right]^{-1}(\mathbf{X}\mathbf{G})'\mathbf{y} = \left[\mathbf{G}'\mathbf{X}'\mathbf{X}\mathbf{G}\right]^{-1}\mathbf{G}'\mathbf{X}'\mathbf{y} \\
&= \mathbf{G}^{-1}(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{G}')^{-1}\mathbf{G}'\mathbf{X}'\mathbf{y} = \mathbf{G}^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\
&= \mathbf{G}^{-1}\hat{\beta}.
\end{aligned}$$

(b)

$$\begin{aligned}
Var[\hat{\gamma}] &= E\left[(\hat{\gamma}-E[\hat{\gamma}])(\hat{\gamma}-E[\hat{\gamma}])'\right] = E\left[\left(\mathbf{G}^{-1}\hat{\beta}-E\left[\mathbf{G}^{-1}\hat{\beta}\right]\right)\left(\mathbf{G}^{-1}\hat{\beta}-E\left[\mathbf{G}^{-1}\hat{\beta}\right]\right)'\right] \\
&= E\left[\mathbf{G}^{-1}\left(\hat{\beta}-E\left[\hat{\beta}\right]\right)\left(\mathbf{G}^{-1}\left(\hat{\beta}-E\left[\hat{\beta}\right]\right)\right)'\right] = E\left[\mathbf{G}^{-1}\left(\hat{\beta}-E\left[\hat{\beta}\right]\right)\left(\hat{\beta}-E\left[\hat{\beta}\right]\right)'\left(\mathbf{G}^{-1}\right)'\right] \\
&= \mathbf{G}^{-1}E\left[\left(\hat{\beta}-E\left[\hat{\beta}\right]\right)\left(\hat{\beta}-E\left[\hat{\beta}\right]\right)'\right]\left(\mathbf{G}^{-1}\right)' \\
&= \mathbf{G}^{-1}Var\left[\hat{\beta}\right]\left(\mathbf{G}^{-1}\right)'.
\end{aligned}$$

(c) In the regression of $\mathbf{y}$ on $\mathbf{Z}$ we have

$$\begin{aligned}
s_Z^2 &= \frac{\mathbf{e}_Z'\mathbf{e}_Z}{n-k} = \frac{\varepsilon'\mathbf{M}_Z\varepsilon}{n-k} = \frac{\varepsilon'\left(\mathbf{I}_n - \mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\right)\varepsilon}{n-k} \\
&= \frac{\varepsilon'\left(\mathbf{I}_n - \mathbf{XG}\left(\mathbf{G}'\mathbf{X}'\mathbf{XG}\right)^{-1}\left(\mathbf{XG}\right)'\right)\varepsilon}{n-k} = \frac{\varepsilon'\left(\mathbf{I}_n - \mathbf{XG}\left(\mathbf{G}'\mathbf{X}'\mathbf{XG}\right)^{-1}\left(\mathbf{XG}\right)'\right)\varepsilon}{n-k} \\
&= \frac{\varepsilon'\left(\mathbf{I}_n - \mathbf{XGG}^{-1}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\left(\mathbf{G}'\right)^{-1}\mathbf{G}'\mathbf{X}'\right)\varepsilon}{n-k} = \frac{\varepsilon'\left(\mathbf{I}_n - \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\right)\varepsilon}{n-k} \\
&= \frac{\varepsilon'\mathbf{M}_X\varepsilon}{n-k} = \frac{\mathbf{e}_X'\mathbf{e}_X}{n-k} = s_X^2.
\end{aligned}$$

(d) In this case, $\mathbf{G} = 10\mathbf{I}_k$. We know that $\hat{\gamma} = \mathbf{G}^{-1}\hat{\beta} = \left(10\mathbf{I}_k\right)^{-1}\hat{\beta} = \frac{1}{10}\hat{\beta}$. Furthermore,

$$Var[\hat{\gamma}] = \mathbf{G}^{-1}Var\left[\hat{\beta}\right]\left(\mathbf{G}^{-1}\right)' = \left(10\mathbf{I}_k\right)^{-1}Var\left[\hat{\beta}\right]\left(10\mathbf{I}_k\right)^{-1} = \frac{1}{100}Var\left[\hat{\beta}\right] = \frac{\sigma^2}{100}\left(\mathbf{X}'\mathbf{X}\right)^{-1}.$$

Of course we estimate σ^2 using $s_Z^2 = s_X^2$, and obtain

$$\widehat{Var}[\hat{\gamma}] = \frac{s_X^2}{100}\left(\mathbf{X}'\mathbf{X}\right)^{-1} = \frac{1}{100}\widehat{Var}\left[\hat{\beta}\right].$$

We construct the t-test statistic

$$t = \frac{\hat{\gamma}}{\sqrt{\widehat{Var}[\hat{\gamma}]}} = \frac{\frac{1}{10}\hat{\beta}}{\sqrt{\frac{1}{100}\widehat{Var}\left[\hat{\beta}\right]}} = \frac{\frac{1}{10}\hat{\beta}}{\frac{1}{10}\sqrt{\widehat{Var}\left[\hat{\beta}\right]}} = \frac{\hat{\beta}}{\sqrt{\widehat{Var}\left[\hat{\beta}\right]}}$$

which is the same test statistic we would use if we had regressed $\mathbf{y}$ on $\mathbf{X}$.

4. The various estimators of β_2 are (most of these make use of the fact that $\mathbf{M}_1$ and $\mathbf{P}_1$ are idempotent):

(a) $\hat{\beta}_2 = \left(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2\right)^{-1}\mathbf{X}_2\mathbf{M}_1\mathbf{y}$ (from the partitioned inverse formula, or directly from solving the least squares normal equations for β_2).

(b) $\hat{\beta}_2 = \left(\mathbf{X}_2'\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{P}_1\mathbf{y}$ (from the usual formula)

(c) $\hat{\beta}_2 = \left(\mathbf{X}_2'\mathbf{P}_1\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{P}_1\mathbf{y}$ (from the usual formula)

(d) $\hat{\beta}_2 = \left(\mathbf{X}_2'\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y}$ (from the usual formula)

(e) $\hat{\beta}_2 = \left(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y}$ (from the usual formula)

(f) $\hat{\beta}_2 = \left(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y}$ (from the usual formula)

(g) $\hat{\beta}_2 = \left(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y}$ (from the partitioned inverse formula, or directly from solving the least squares normal equations for β_2)

(h) $\hat{\beta}_2 = \left(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y}$ (from the usual formula, noting that $\mathbf{M}_1\mathbf{X}_1 = \mathbf{0}$).

Hence there are four different estimators of β_2: (a), (e), (f), (g), (h) are all the same, and (b), (c), (d) are all different.