Econometrics 710

Midterm Exam, Sample Answers

March 4, 1999

1. 15 points

   (a) $\tilde{\beta} = (Z'Z)^{-1} Z'Y = (B'X'XB)^{-1} B'X'Y = \left(B^{-1}(X'X)^{-1}(B')^{-1}\right) B'X'Y = B^{-1}(X'X)^{-1} X'Y = B^{-1}\hat{\beta}$

   (b) $\tilde{e} = Y - Z\tilde{\beta} = Y - XBB^{-1}\hat{\beta} = Y - X\hat{\beta} = \hat{e}$

2. 15 points

   It might be helpful to observe that since X has full rank, $X'X > 0$, and since $\lambda > 0$, $\lambda I_k > 0$, thus $X'X + \lambda I_k > 0$. This means that $\hat{\beta}$ exists as defined.

   (a) Let $e = Y - E(Y \mid X) = Y - X\beta$ so that $E(e \mid X) = 0$. Then

   $$\hat{\beta} = (X'X + \lambda I_k)^{-1} X'(X\beta + e) = (X'X + \lambda I_k)^{-1} X'X\beta + (X'X + \lambda I_k)^{-1} X'e.$$

   and

   $$\begin{aligned} E\left(\hat{\beta} \mid X\right) &= (X'X + \lambda I_k)^{-1} X'X\beta + (X'X + \lambda I_k)^{-1} X'E(e \mid X) \\ &= (X'X + \lambda I_k)^{-1} X'X\beta \\ &= (X'X + \lambda I_k)^{-1} (X'X + \lambda I_k - \lambda I_k)\beta \\ &= (X'X + \lambda I_k)^{-1} (X'X + \lambda I_k)\beta + (X'X + \lambda I_k)^{-1} (-\lambda I_k)\beta \\ &= \beta - (X'X + \lambda I_k)^{-1} \beta\lambda. \end{aligned}$$

   (b) The conditional bias is $b = E\left(\hat{\beta} - \beta \mid X\right) = -(X'X + \lambda I_k)^{-1} \beta\lambda$. Since $(X'X + \lambda I_k)^{-1} > 0$, there is no vector $\alpha \neq 0$ such that $(X'X + \lambda I_k)^{-1} \alpha = 0$. Therefore $b \neq 0$ if $\beta\lambda \neq 0$. Note that $b = 0$ if $\beta = 0$. Thus $\hat{\beta}$ is biased unless $\beta = 0$.

3. 30 points

   Let $e = Y^* - E(Y^* \mid X)$. Then

   $$\begin{aligned} Y &= Y^* + u \\ &= E(Y^* \mid X) + e + u \\ &= X\beta + v, \end{aligned} \tag{1}$$

where $v = e + u$. Note that

$$\begin{aligned} E(v \mid X) &= E(e \mid X) + E(u \mid X) \\ &= 0 + E(E(u \mid Y^*, X) \mid X) \\ &= 0, \end{aligned} \tag{2}$$

which shows that (1) is a valid regression model. This means that all of the analysis for the traditional regression model applies, only that the equation error is v, rather than the ideal error e.

(a) $E(Y \mid X) = E(Y^* + u \mid X) = E(Y^* \mid X) + E(u \mid X) = X\beta$.

(b) Equations (1) and (2) show that the conditions for OLS estimation are satisified. Thus $\hat{\beta} \to_p \beta$ if the moment conditions are satisfied ($E|x_i|^2 < \infty$ and $E|v_i|^2 < \infty$) and the design matrix is full rank ($Ex_i x_i' > 0$).

(c) Similarly, $\sqrt{n}\left(\hat{\beta} - \beta\right) \to_d N(0, V)$, where $V = (Ex_i x_i')^{-1} E(x_i x_i' v_i^2)(Ex_i x_i')^{-1}$. This is a fine answer. A complete answer would go a bit further. Observe that since the information in (y_i^*, x_i) and (e_i, x_i) is the same, then

$$E(u_i \mid y_i^*, x_i) = E(u_i \mid e_i, x_i) = 0.$$

Hence

$$E(e_i u_i \mid x_i) = E(e_i E(u_i \mid e_i, x_i) \mid x_i) = 0,$$

so e_i and u_i are uncorrelated (conditional on x_i). Thus since $v_i^2 = e_i^2 + 2e_i u_i + u_i^2$, we see that

$$\begin{aligned} E\left(x_i x_i' v_i^2\right) &= E\left(x_i x_i' e_i^2\right) + E(x_i x_i' e_i u_i) + E\left(x_i x_i' u_i^2\right) \\ &= E\left(x_i x_i' e_i^2\right) + E\left(x_i x_i' u_i^2\right). \end{aligned}$$

Hence the asymptotic variance of $\hat{\beta}$ can be written as

$$V = (Ex_i x_i')^{-1}\left(E\left(x_i x_i' e_i^2\right) + E\left(x_i x_i' u_i^2\right)\right)(Ex_i x_i')^{-1}.$$

Note that $V > V_0 = (Ex_i x_i')^{-1} E(x_i x_i' e_i^2)(Ex_i x_i')^{-1}$, the asymptotic variance obtained from the OLS regression of Y^* on X.

4. 30 points

(a) Indeed, the variables are highly correlated. $\hat{\beta}_1$ and $\hat{\beta}_2$ have a correlation of $-.5/.7 \simeq .71$. In this context, individual coefficients are less precisely estimated than when the correlation is closer to zero. We observe in this example that the standard errors appear quite large (relative to the magnitude of the coefficient estimates).

(b) The regression of profits on sales takes the form $x_2 = \hat{\alpha} x_1 + \hat{x}_2^*$, where $\hat{\alpha} = \frac{x_1' x_2}{x_1' x_1} = \frac{8}{10} = .8$. Hence

$$x_2^* = x_2 - .8x_1.$$

(c) As shown in problem 1, if X is replaced by a set of linear combinations, the OLS coefficients are similarly recombined. In the present case, all you need to note is that $x_2 = x_2^* + .8x_1$, so

$$\begin{aligned} \hat{y} &= .5x_1 + .4x_2 \\ &= .5x_1 + .4\left(x_2^* + .8x_1\right) \\ &= \left(.5 + (.4)(.8)\right) x_1 + .4x_2^* \\ &= .82x_1 + .4x_2^* \end{aligned}$$

Hence

$$\begin{aligned} \tilde{\beta}_1 &= .82 \\ \tilde{\beta}_2 &= .4. \end{aligned}$$

To do this calculation using the notation of problem 1, let $Z = [x_1\ x_2^*]$. Then $Z = XB$ if

$$B = \begin{bmatrix} 1 & -.8 \\ 0 & 1 \end{bmatrix}.$$

Note that

$$B^{-1} = \begin{bmatrix} 1 & .8 \\ 0 & 1 \end{bmatrix}.$$

From problem 1, $\tilde{\beta} = B^{-1}\hat{\beta}$, so $\tilde{\beta}_1 = \tilde{\beta}_1 + .8\tilde{\beta}_2$ and $\tilde{\beta}_2 = \hat{\beta}_2$ as we found above.

(d) Since $\tilde{\beta} = B^{-1}\hat{\beta}$ and $\hat{\beta}$ has covariance matrix V, the covariance matrix for $\tilde{\beta}$ is $B^{-1}VB^{-1\prime}$, which is

$$\begin{bmatrix} 1 & .8 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} .7 & -.5 \\ -.5 & .7 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ .8 & 1 \end{bmatrix} = \begin{bmatrix} .348 & .06 \\ .06 & .7 \end{bmatrix}.$$

The strange thing is that this doesn't make sense, because the regressors are orthogonal, so the off-diagonal terms of the covariance matrix should be zero. The reason for the discrepancy is that our reporting of the covariance matrix made a rounding simplification. A consistent result is obtained if the covariance matrix of $\hat{\beta}$ is $\begin{bmatrix} .7 & -.56 \\ -.56 & .7 \end{bmatrix}$, in which case the covariance matrix for $\tilde{\beta}$ is

$$\begin{bmatrix} 1 & .8 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} .7 & -.56 \\ -.56 & .7 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ .8 & 1 \end{bmatrix} = \begin{bmatrix} .252 & 0 \\ 0 & .7 \end{bmatrix}.$$

Another way to calculate the standard errors is to note that $\tilde{\beta}_1 = \begin{pmatrix} 1 & .8 \end{pmatrix} \hat{\beta}$ so has variance $\begin{pmatrix} 1 & .8 \end{pmatrix} \begin{bmatrix} .7 & -.5 \\ -.5 & .7 \end{bmatrix} \begin{pmatrix} 1 \\ .8 \end{pmatrix} = .348$, and $\tilde{\beta}_2 = \hat{\beta}_1$ so has variance .7. The standard errors are the square roots of the diagonals, so

$$
\begin{aligned}
s\left(\tilde{\beta}_1\right) &= \sqrt{.348} \simeq .59 \\
s\left(\tilde{\beta}_2\right) &= \sqrt{.7} \simeq .83.
\end{aligned}
$$

(e) It is impossible to affect collinearity. Collinearity is a property of the data, and if the coefficients of interest are imprecisely estimated, that's a fact that cannot be made to go away. On the other hand, note that the proposed rotation of the data results in coefficient estimates which are nearly uncorrelated, and as a result one of the estimates is more precisely estimated (it has a smaller standard error). The silliness of the proposal, however, is that this particular parameterization is *ad hoc* and is unlikely to be of particular interest to the researcher.