PLSC 503: Problem Set 2 Solutions

Thad Dunning

Department of Political Science

Yale University

Spring 2010

1 Theoretical/Conceptual Problems

Question 1: What is the size of $X'X$? Of $X'Y$? What about $(X'X)^{-1}$ and $(X'X)^{-1}(X'Y)$? Can you multiply X and Y? Why or why not?

Solution. The matrix X is $n \times 2$, so X' is $2 \times n$ and $X'X$ is a 2×2 square matrix. Since Y is $n \times 1$, $X'Y$ is 2×1. The square matrix $(X'X)^{-1}$ is 2×2. Combining the previous results, we have that $(X'X)^{-1}(X'Y)$ is a 2×1 column vector. Finally, since X is $n \times 2$ and Y is $n \times 1$, X and Y are not conformable: you cannot multiply X and Y.

Question 2: Find $X'X$.

Solution: Following the usual rules for matrix multiplication gives

$$X'X = \begin{pmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} (x_i)^2 \end{pmatrix} \tag{1}$$

Question 3: Find $(X'X)^{-1}$.

Solution:

The first step in finding the inverse is to find the determinant of $(X'X)$ (see Freedman 2007, Chapter 3). Using the general rule for the determinant of a 2×2 matrix, this is

$$\det(X'X) = n \sum_{i=1}^{n} (x_i)^2 - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} x_i). \tag{2}$$

Next, we need the adjoint of $(X'X)$. Remember that the adjoint is the transpose of the matrix of cofactors. The matrix of cofactors of $X'X$ is

$$\begin{pmatrix} \sum_{i=1}^{n} (x_i)^2 & -\sum_{i=1}^{n} x_i \\ -\sum_{i=1}^{n} x_i & n \end{pmatrix} \tag{3}$$

and the adjoint is the transpose of this matrix:

$$\begin{pmatrix} \sum_{i=1}^{n}(x_i)^2 & -\sum_{i=1}^{n} x_i \\ -\sum_{i=1}^{n} x_i & n \end{pmatrix} \tag{4}$$

(In this case, the matrix of cofactors is symmetric, so the matrix of cofactors and its transpose are the same matrix.)

Thus, since $A^{-1} = \text{adj}(A)/\det(A)$, we then have

$$(X'X)^{-1} = \frac{1}{n\sum_{i=1}^{n}(x_i)^2 - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} x_i)} \begin{pmatrix} \sum_{i=1}^{n}(x_i)^2 & -\sum_{i=1}^{n} x_i \\ -\sum_{i=1}^{n} x_i & n \end{pmatrix} \tag{5}$$

To verify that $(X'X)(X'X)^{-1} = I_{2\times 2}$, note that

$$(X'X)(X'X)^{-1} = \frac{1}{\det(X'X)} \begin{pmatrix} n\sum_{i=1}^{n}(x_i)^2 - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} x_i) & -n\sum_{i=1}^{n} x_i + n\sum_{i=1}^{n} x_i \\ (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i^2)(\sum_{i=1}^{n} x_i) & -(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} x_i) + n\sum_{i=1}^{n}(x_i)^2 \end{pmatrix} \tag{6}$$

where $\det(X'X) = n\sum_{i=1}^{n}(x_i)^2 - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} x_i)$. Note the off-diagonal elements reduce to 0, while the division by $\det(X'X)$ reduces the diagonal elements to 1, and so

$$(X'X)(X'X)^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{7}$$

Question 4: Find $(X'Y)$.

Solution:.

$$(X'Y) = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{pmatrix} \tag{8}$$

Question 5: Find $(X'X)^{-1}X'Y$.

Solution:. Post-multiplying (5) by (8), we have

$$(X'X)^{-1}X'Y = \begin{pmatrix} \frac{\sum_{i=1}^{n}(x_i^2)\sum_{i=1}^{n}(y_i)-(\sum_{i=1}^{n}x_i)(\sum_{i=1}^{n}x_iy_i)}{n\sum_{i=1}^{n}(x_i)^2-(\sum_{i=1}^{n}x_i)(\sum_{i=1}^{n}x_i)} \\ \\ \frac{-(\sum_{i=1}^{n}x_i)(\sum_{i=1}^{n}y_i)+n\sum_{i=1}^{n}x_iy_i}{n\sum_{i=1}^{n}(x_i)^2-(\sum_{i=1}^{n}x_i)(\sum_{i=1}^{n}x_i)} \end{pmatrix} \tag{9}$$

Question 6: Show that the $(2,1)$ element of $(X'X)^{-1}X'Y = \frac{\text{Cov}(x,y)}{\text{Var}(x)}$.

Solution:

Note from (9) that the $(2,1)$ element of $(X'X)^{-1}X'Y$ can be written as

$$\frac{n\sum_{i=1}^{n}x_iy_i - (\sum_{i=1}^{n}x_i)(\sum_{i=1}^{n}y_i)}{n\sum_{i=1}^{n}(x_i)^2 - (\sum_{i=1}^{n}x_i)(\sum_{i=1}^{n}x_i)}. \tag{10}$$

Now divide both the numerator and the denominator of (10) by n^2 (that is, multiply through by $\frac{n^2}{n^2}$)

to get

$$\frac{\overline{xy} - (\bar{x})(\bar{y})}{\overline{x^2} - \bar{x}^2} = \frac{\text{Cov}(x,y)}{\text{Var}(x)}, \tag{11}$$

where the equality follows from the alternate definitions of covariance and variance.

Question 7: Show that the $(1,1)$ element of $(X'X)^{-1}X'Y = \bar{y} - b\bar{x}$, where $b = \frac{\text{Cov}(x,y)}{\text{Var}(x)}$.

Solution:

Note first that $\bar{y} - b\bar{x}$ can be written as

$$(\bar{y} - \frac{\text{Cov}(x,y)}{\text{Var}(x)}\bar{x}) = \frac{\bar{y}\text{Var}(x) - \text{Cov}(x,y)\bar{x}}{\text{Var}(x)}.(12)$$

Thus, we just need to show that equation (12) is equal to the $(1,1)$ element of $(X'X)^{-1}X'Y$ in (9). I first present the proof, then explain the steps:

$$\frac{\sum_{i=1}^{n}(x_i^2)\sum_{i=1}^{n}(y_i) - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} x_i y_i)}{n\sum_{i=1}^{n}(x_i)^2 - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} x_i)} = \frac{(\overline{x^2})(\bar{y}) - (\bar{x})(\overline{xy})}{\overline{x^2} - \bar{x}^2} \tag{13}$$

$$= \frac{(\overline{x^2})(\bar{y}) - \bar{y}\bar{x}^2 - (\bar{x})(\overline{xy}) + \bar{y}\bar{x}^2}{\overline{x^2} - \bar{x}^2} \tag{14}$$

$$= \frac{\bar{y}[\overline{x^2} - \bar{x}^2] - \bar{x}[\overline{xy} - \bar{y}\bar{x}]}{\overline{x^2} - \bar{x}^2} \tag{15}$$

$$= \frac{\bar{y}\mathrm{Var}(x) - \mathrm{Cov}(x,y)\bar{x}}{\mathrm{Var}(x)}. \tag{16}$$

To get to the right-hand side of (13), we multiply by $\frac{n^2}{n^2}$. Next, we add and subtract $\bar{y}\bar{x}^2$ to the numerator (second line) and factor terms (third lines). Finally, we use the alternate definitions of covariance and variance, which completes the proof.

Note that in problems 1-7, we are doing bivariate regression. The first column of X is a column of 1s, the second column is the vector of x_is. The 2×1 vector $(X'X)^{-1}X'Y$ is the "OLS estimator." The $(2, 1)$ element of this vector gives the slope coefficient for the bivariate regression that we derived previously. The $(1, 1)$ element is the intercept. The point of the problem is to show how the matrix representation generates the usual formulas for the fitted intercept and slope coefficient, in a bivariate regression.

Question 8: Freedman (2009), exercise 3.B.12

If u and v are $n \times 1$, show that $||u + v||^2 = ||u||^2 + ||v||^2 + 2u \cdot v$.

Solution: First, note that the $n \times 1$ vector $(u + v)$ has typical element $(u_i + v_i)$, so the inner product

$(u + v) \cdot (u + v)$ has typical element $(u_i + v_i)^2$. In fact,

$$\begin{aligned}
(u+v)\cdot(u+v) &= \sum_{i=1}^{n}(u_i+v_i)^2 \\
&= \sum_{i=1}^{n}(u_i^2+2u_iv_i+v_i^2) \\
&= \sum_{i=1}^{n}u_i^2+2\sum_{i=1}^{n}(u_iv_i)+\sum_{i=1}^{n}v_i^2) \\
&= ||u||^2+2u\cdot v+||v||^2 \qquad (17)
\end{aligned}$$

In (17), the second line follows from writing out terms, while the third follows from distributing the summation sign and the fourth follows from the definition of the inner product.

Question 9: Freedman (2009), exercise 3.B.13

If u and v are $n \times 1$, show that $||u + v||^2 = ||u||^2 + ||v||^2$ if and only if $u \perp v$.

Solution: The "if" direction follows from the final line of (17): if $u \perp v$, $2u \cdot v = 0$, so $(u + v) \cdot (u + v) = ||u||^2 + ||v||^2$. The "only if" direction is similar: by (17), $||u + v||^2 = ||u||^2 + ||v||^2 + 2u \cdot v$, and this equal to zero only if $u \cdot v = 0$. But that's the definition of orthogonality, so $||u + v||^2 = ||u||^2 + ||v||^2$ only if $u \perp v$.

Comment: This is Pythagoras' theorem in n dimensions. The idea plays a key role in the theory of regression, since the vector of regression residuals e is always orthogonal to each column of X. So, for example, if $Y = X\hat{\beta} + e$, we have that $||Y||^2 = ||X\hat{\beta} + e||^2 = ||X\hat{\beta}||^2 + ||e||^2$.

Question 10: Freedman (2009), exercise 3.B.14 Suppose X is $n \times p$ with rank $p < n$. Suppose Y is $n \times 1$. Let $\hat{\beta} = (X'X)^{-1}X'Y$ and $e = Y - X\hat{\beta}$.

(a) **Question:** Show that $X'X$ is $p \times p$, while $X'Y$ is $p \times 1$.

Solution: X is $n \times p$, so X' is $p \times n$, and $X'X$ is $p \times p$. Y is $n \times 1$, so $X'Y$ is $p \times 1$.

(b) **Question:** Show that $X'X$ is symmetric.

Solution: Following the hint, exercise 3.B.9(a) in Freedman (2009) shows that $(MN)' = N'M'$, so $(X'X)' = X'(X')' = X'X$. Thus, by the definition of symmetry (a matrix A is symmetric if $A = A'$), $X'X$ is symmetric.

(c) **Question:** Show that $X'X$ is invertible.

Solution: Exercise 3.B.10 in Freedman (2009) shows that $X'X$ has rank p, since X has rank p. Thus, we are done: $X'X$ has full rank, so it is invertible.

A digression on exercise 3.B.10: Following the hints in that exercise, suppose X has rank p and c is $p \times 1$. Then $X'Xc = 0_{p\times 1} \Rightarrow c'X'Xc = 0 \Rightarrow \|Xc\|^2 \Rightarrow Xc = 0_{n\times 1} \Rightarrow c = 0_{n\times 1}$. The key is that we assumed X has rank p, so if $Xc = 0_{n\times 1}$, then $c = 0_{n\times 1}$; otherwise, one of the columns of X can be written as a linear combination of the others, and thus X does not have full rank of p, which contradicts the assumption.

Thus, $X'Xc = 0_{p\times 1}$ if and only if $c = 0_{n\times 1}$, which implies that $X'X$ has full rank.

Try it for a 2×2 matrix. E.g., if

$$X'X = \begin{pmatrix} w & x \\ y & z \end{pmatrix} \quad \text{and} \quad c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \tag{18}$$

then $X'Xc = 0_{2\times 1}$ implies

$$X'Xc = \begin{pmatrix} w & x \\ y & z \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} w \times c_1 + x \times c_2 \\ y \times c_1 + z \times c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{19}$$

That is,

$$\begin{aligned} w \times c_1 + x \times c_2 &= 0 \\ y \times c_1 + z \times c_2 &= 0, \end{aligned} \tag{20}$$

which can be rewritten as

$$\begin{aligned} w \times c_1 &= -x \times c_2 \\ y \times c_1 &= -z \times c_2. \end{aligned}$$

or

$$\begin{aligned} w \times \frac{-c_1}{c_2} &= x \\ y \times \frac{-c_1}{c_2} &= z. \end{aligned}$$

This implies that we can write one column of $X'X$ as a linear combination of the other column:

$$\begin{pmatrix} w \\ y \end{pmatrix} \frac{-c_1}{c_2} = \begin{pmatrix} x \\ z \end{pmatrix}. \tag{21}$$

(d) **Question:** Show that $(X'X)^{-1}$ is $p \times p$, so $\hat{\beta} = (X'X)^{-1}X'Y$ is $p \times 1$.

Solution: $X'X$ is $p \times p$ – see problem (a) – and has an inverse – see problem (c) – so we must have $(X'X)(X'X)^{-1} = I_{p \times p}$; thus, $(X'X)^{-1}$ is $p \times p$. Since $X'Y$ is $p \times 1$ by problem (a), $(X'X)^{-1}X'Y$ is $p \times 1$.

(e) **Question:** Show that $(X'X)^{-1}$ is symmetric.

Solution: By 3.B.9(b), if the square matrices M and N are both invertible, $(M')^{-1} = (M^{-1})'$. We want to show that $(X'X)^{-1}$ is symmetric, that is, $((X'X)^{-1})' = (X'X)^{-1}$. Just apply the rule:

$((X'X)^{-1})' = ((X'X)')^{-1}$, and $(X'X)' = X'X$ by the symmetry of $X'X$ – see problem (b).

(f) **Question:** Show that $X\hat{\beta}$ and $e = Y - X\hat{\beta}$ are $n \times 1$.

Solution: We showed in (d) that $\hat{\beta}$ is $p \times 1$, so $X\hat{\beta}$ is $n \times 1$. So is Y, thus, so is $e = Y - X\hat{\beta}$.

(g) **Question:** Show that $X'X\hat{\beta} = X'Y$, and hence $X'e = 0_{p\times 1}$.

Solution: For the first part, just substitute for $\hat{\beta}$: $X'X\hat{\beta}$ =X'X(X'X)$^{-1}X'Y = I_{p\times p}X'Y = X'Y$, where the penultimate equality follows by the definition of the inverse. For the second part, substitute for e and multiply: $X'e = X'(Y - X\hat{\beta}) = X'Y - X'X'\hat{\beta} = X'Y - X'Y = 0_{p\times 1}$.

(h) **Question:** Show that $e \perp X\hat{\beta}$, so $||Y||^2 = ||X\beta||^2 + ||e||^2$.

Solution: First, $e'(X\hat{\beta}) = (e'X)\hat{\beta} = (X'e)'\hat{\beta} = 0$, because we showed in (g) that $X'e = 0_{p\times 1}$. Thus, $e \perp X\hat{\beta}$. The second part follows from exercises 3.B.12 and 3.B.13: if $Y = X\beta + e$, we have that $||Y||^2 = ||X\beta + e||^2 = ||X\beta||^2 + ||e||^2 + 2e \cdot (X\hat{\beta}) = ||X\beta||^2 + ||e||^2$ by the orthogonality of e and $X\hat{\beta}$.

(i) **Question:** If γ is $p \times 1$, show that $||Y - X\gamma||^2 = ||Y - X\hat{\beta}||^2 + ||X(\hat{\beta} - \gamma)||^2$

Solution: From the hint, just add and subtract $X\hat{\beta}$ to $Y - X\gamma$: that is, $Y - X\gamma = Y - X\hat{\beta} + X(\hat{\beta} - \gamma)$. Now, $Y - X\hat{\beta}$ is an $n \times 1$ vector and so is $X(\hat{\beta} - \gamma)$. Thus, by exercise 3.B.12, $||Y - X\hat{\beta} + X(\hat{\beta} - \gamma)||^2 = ||Y - X\hat{\beta}||^2 + ||X(\hat{\beta} - \gamma)||^2 + 2(Y - X\hat{\beta}) \cdot X(\hat{\beta} - \gamma)$. You can show that the last term goes to zero, that is, $Y - X\hat{\beta} \perp X(\hat{\beta} - \gamma)$. (Try it!).

(j) **Question:** Show that $||Y - X\gamma||^2$ is minimized when $\gamma = \hat{\beta}$.

Solution: By (i), $||Y - X\gamma||^2 = ||Y - X\hat{\beta}||^2 + ||X(\hat{\beta} - \gamma)||^2$. The first term on the right-hand side is at least zero. We minimize by the expression by sending the second term on the right-had side to zero: this happens if and only if $\gamma = \hat{\beta}$.

(j) **Question:** If $\tilde{\beta}$ is $p \times 1$ with $Y - X\tilde{\beta} \perp X$, show that $\tilde{\beta} = \hat{\beta}$.

Solution: $X'(Y - X\tilde{\beta}) = X'Y - X'X\tilde{\beta} = 0$, by the orthogonality of X and $Y - X\tilde{\beta}$. Now, $X'(Y - X\hat{\beta}) = X'Y - X'X\hat{\beta} = 0$, since $X'X\hat{\beta} = X'X(X'X)^{-1}X'Y = X'Y$. Thus, we must have $X'X\tilde{\beta} = X'X\hat{\beta}$, so $\tilde{\beta} = \hat{\beta}$.

Note: answers to 3.B.14 (l) and (m) can be found in the Answers to Exercises section in Freedman (2009).

Question 11: Freedman (2009), exercise 3.B.17.

Solution: The solution can be found in the Answers to Exercises section in Freedman (2009). Note that Angrist and Pischke (2009: 35) give an alternate statement of this "Frisch-Waugh" theorem, which they call the "regression anatomy" formula.

2 Computer lab: Yule's regression

A stata .do file solving the computer exercises is posted on under the "Resources" tab of the classes v2 server, courtesy of Mario Chacón.