Economics 671  
Solutions: Problem Set #1  

(1) First, let’s expand the traditional formula. Note that

\[ \sum_{i}(x_i - \overline{x})(y_i - \overline{y}) = \sum_{i}x_iy_i - n\overline{x}\overline{y} = \sum_{i}x_iy_i - \frac{1}{n}(\sum_{i}x_i)(\sum_{i}y_i). \]

Likewise, for the denominator:

\[ \sum_{i}(x_i - \overline{x})^2 = \sum_{i}x_i^2 - n\overline{x}^2 = \sum_{i}x_i^2 - \frac{1}{n}(\sum_{i}x_i)^2. \]

Multiplying each of these by \( n \), it follows that the traditional formula for the OLS slope estimator is equivalent to:

\[ \hat{\beta}_2 = \frac{n\sum_{i}x_iy_i - (\sum_{i}x_i)(\sum_{i}y_i)}{n\sum_{i}x_i^2 - (\sum_{i}x_i)^2}. \]

Now, direct application of our \((X'X)^{-1}X'y\) formula gives:

\[ \hat{\beta} = \left[ \frac{n}{\sum_{i}x_i^2} \sum_{i}x_i \sum_{i}x_i^2 \right]^{-1} \left[ \sum_{i}y_i \sum_{i}x_i^2 \overline{y} \right]. \]

Inverting the \( 2 \times 2 \) matrix and picking of the second element of \( \hat{\beta} \) shows that

\[ \hat{\beta}_2 = \frac{n\sum_{i}x_iy_i - (\sum_{i}x_i)(\sum_{i}y_i)}{n\sum_{i}x_i^2 - (\sum_{i}x_i)^2}, \]

which is the same as the above.

(2a) We would not expect the full rank condition to be violated, as with even a small typical sample, we would not expect an exact linear relationship between \( \text{Age} \) and \( \text{Education} \). However, there are reasons to think that \( E(u | \text{Age}, \text{Ed}) \not= 0 \). Just as an example: if unobserved “ability” is excluded from the regression model, we would expect that ability is an independent determinant of wages \( (y) \) and is also correlated with educational attainment.

(2b) Consider equation (2):

\[ \text{Wage}_i = \theta_0 + \theta_1 \text{Ed}_i + \theta_2 \text{Exper}_i + \epsilon_i \]

\[ = \theta_0 + \theta_1 \text{Ed}_i + \theta_2 (\text{Age}_i - \text{Ed}_i - 6) + \epsilon_i \]

\[ = (\theta_0 - 6\theta_2) + \theta_2 \text{Age}_i + (\theta_1 - \theta_2) \text{Ed}_i + \epsilon_i \]

This is in the same form as equation (1), with the following relationships among the coefficients:

\[ \beta_0 = (\theta_0 - 6\theta_2), \quad \beta_1 = \theta_2, \quad \beta_2 = (\theta_1 - \theta_2). \]
We would expect the coefficient estimates to be related in exactly the way described above. We will now prove this.

(2c) First, define the following matrices, noting that each represent the stacked covariate matrices in equations (1) and (2):

\[
X_{n \times 3} = \begin{bmatrix}
1 & \text{Age}_1 & \text{Ed}_1 \\
1 & \text{Age}_2 & \text{Ed}_2 \\
\vdots & \ddots & \vdots \\
1 & \text{Age}_n & \text{Ed}_n
\end{bmatrix},
Z_{n \times 3} = \begin{bmatrix}
1 & \text{Ed}_1 & \text{Exper}_1 \\
1 & \text{Ed}_2 & \text{Exper}_2 \\
\vdots & \ddots & \vdots \\
1 & \text{Ed}_n & \text{Exper}_n
\end{bmatrix}.
\]

In addition, let us define a $3 \times 3$, invertible matrix $A$ as follows

\[
A = \begin{bmatrix}
1 & 0 & -6 \\
0 & 0 & 1 \\
0 & 1 & -1
\end{bmatrix}.
\]

Given this construction, we can see that

\[
XA = \begin{bmatrix}
1 & \text{Age}_1 & \text{Ed}_1 \\
1 & \text{Age}_2 & \text{Ed}_2 \\
\vdots & \ddots & \vdots \\
1 & \text{Age}_n & \text{Ed}_n
\end{bmatrix} \begin{bmatrix}
1 & 0 & -6 \\
0 & 0 & 1 \\
0 & 1 & -1
\end{bmatrix} = \begin{bmatrix}
1 & \text{Ed}_1 & \text{Age}_1 - \text{Ed}_1 - 6 \\
1 & \text{Ed}_2 & \text{Age}_2 - \text{Ed}_2 - 6 \\
\vdots & \ddots & \ddots \\
1 & \text{Ed}_n & \text{Age}_n - \text{Ed}_n - 6
\end{bmatrix} = \begin{bmatrix}
1 & \text{Ed}_1 & \text{Exper}_1 \\
1 & \text{Ed}_2 & \text{Exper}_2 \\
\vdots & \ddots & \ddots \\
1 & \text{Ed}_n & \text{Exper}_n
\end{bmatrix} = Z.
\]

Equivalently,

\[X = ZA^{-1}.
\]

So, now let us consider the OLS estimator of equation (1):

\[
\hat{\beta} = (X'X)^{-1}X'y = \left[(ZA^{-1})(ZA^{-1})\right]^{-1}(ZA^{-1})'y = \left[(A')^{-1}Z'ZA^{-1}\right]^{-1}(A')^{-1}Z'y = A(Z'Z)^{-1}A'(A')^{-1}Z'y = A(Z'Z)^{-1}Z'y = A\tilde{\theta}.
\]

Note that the above uses the following linear algebra properties (assuming the requisite inverses exist): $(AB)' = B'A'$, $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$, $(A^{-1})' = (A')^{-1}$. Written out, the final equation above states

\[
\begin{bmatrix}
\hat{\beta}_0 \\
\hat{\beta}_1 \\
\hat{\beta}_2
\end{bmatrix} = \begin{bmatrix}
1 & 0 & -6 \\
0 & 0 & 1 \\
0 & 1 & -1
\end{bmatrix} \begin{bmatrix}
\hat{\theta}_0 \\
\hat{\theta}_1 \\
\hat{\theta}_2
\end{bmatrix} = \begin{bmatrix}
\hat{\theta}_0 - 6\hat{\theta}_2 \\
\hat{\theta}_1 \\
\hat{\theta}_2 - \hat{\theta}_2
\end{bmatrix}
\]

which is exactly what we established by substituting for Exper, and what we expected would hold among the coefficient estimates as well.