Frisch-Waugh-Lovell Theorem

Chungsang Tom Lam[1]

[1]Department of Economics
Clemson University

January 30, 2019

Partitioned Regression

Suppose the true model is:

$$y = X_1\beta_1 + X_2\beta_2 + \epsilon \tag{1}$$

In other words, we have :

$$y = X\beta + \epsilon$$

$$X = \begin{bmatrix} X_1 & X_2 \end{bmatrix} \text{ and } \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

Recall the solution of OLS $\hat{\beta}$ can be solved from the equation

$$(X'X)\hat{\beta} = X'Y \tag{2}$$

Partitioned Regression

Write the X_1 and X_2 components explicitly,

$$\begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} X_1'y \\ X_2'y \end{bmatrix} \tag{3}$$

We can solve $\hat{\beta}_1$ in terms of $\hat{\beta}_2$,

$$\begin{aligned} (X_1'X_1)\hat{\beta}_1 + (X_1'X_2)\hat{\beta}_2 &= X_1'y \\ (X_1'X_1)\hat{\beta}_1 &= X_1'y - (X_1'X_2)\hat{\beta}_2 \\ \hat{\beta}_1 &= (X_1'X_1)^{-1}X_1'y - (X_1'X_1)^{-1}(X_1'X_2)\hat{\beta}_2 \end{aligned}$$

Partitioned Regression

We can continue to solve for $\hat{\beta}_2$,

$$\begin{aligned}
(X_2'X_1)\hat{\beta}_1 + (X_2'X_2)\hat{\beta}_2 &= X_2'y \\
(X_2'X_1)((X_1'X_1)^{-1}X_1'y - (X_1'X_1)^{-1}(X_1'X_2)\hat{\beta}_2) + (X_2'X_2)\hat{\beta}_2 &= X_2'y \\
(X_2'X_1)(X_1'X_1)^{-1}X_1'y - (X_2'X_1)(X_1'X_1)^{-1}(X_1'X_2)\hat{\beta}_2 + (X_2'X_2)\hat{\beta}_2 &= X_2'y
\end{aligned}$$

$$\begin{aligned}
(X_2'X_2)\hat{\beta}_2 - (X_2'X_1)(X_1'X_1)^{-1}(X_1'X_2)\hat{\beta}_2 &= X_2'y - (X_2'X_1)(X_1'X_1)^{-1}X_1'y \\
((X_2'X_2) - (X_2'X_1)(X_1'X_1)^{-1}(X_1'X_2))\hat{\beta}_2 &= X_2'(I - X_1(X_1'X_1)^{-1}X_1')y \\
X_2'(I - X_1(X_1'X_1)^{-1}X_1')X_2\hat{\beta}_2 &= X_2'(I - X_1(X_1'X_1)^{-1}X_1')y
\end{aligned}$$

Partitioned Regression

Finally,

$$\hat{\beta}_2 = (X_2'(I - X_1(X_1'X_1)^{-1}X_1')X_2)^{-1}X_2'(I - X_1(X_1'X_1)^{-1}X_1')y \quad (4)$$

So what exactly is $I - X_1(X_1'X_1)^{-1}X_1'$?
Consider the residual of the OLS

$$\begin{aligned}\hat{\epsilon} &= y - X\hat{\beta} \\ &= y - X(X'X)^{-1}X'y \\ &= (I - X(X'X)^{-1}X')y\end{aligned}$$

In other words, if we multiply $(I - X(X'X)^{-1}X')$ to y, it produces the residual. We name this matrix $M = I - X(X'X)^{-1}X'$. And similarly, $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$

FWL Theorem

Notice that M matrix is idempotent, which means $MM = M$.

M is a projection matrix, it yields the vector of residuals
another projection matrix is Px, it yields the vector of fitted value

$$\begin{aligned} MM &= (I - X(X'X)^{-1}X')(I - X(X'X)^{-1}X') \\ &= I^2 - 2X(X'X)^{-1}X' + X(X'X)^{-1}X'X(X'X)^{-1}X' \\ &= I - X(X'X)^{-1}X' = M \end{aligned}$$

With the M matrix we can rewrite $\hat{\beta}_2$,

$$\hat{\beta}_2 = (X_2'M_1X_2)^{-1}X_2'M_1y \tag{5}$$

And with the idempotent matrix property,

$$\begin{aligned} \hat{\beta}_2 &= ((M_1X_2)'M_1X_2)^{-1}(M_1X_2)'M_1y \\ &= ((X_2^*)'(X_2^*))^{-1}(X_2^*)'y^* \end{aligned}$$

FWL Theorem

- This formula indicates that there are two ways to solve for the OLS estimate $\hat{\beta}_2$. We can regress y on X, solve for $\hat{\beta}$, and $\hat{\beta}_2$ is inside $\hat{\beta}$.
- Alternatively, we can regress y on X_1, solve for the residual y^*, then regress X_2 on X_1, solve for the residual X_2^*. Regress y^* on X_2^* will yield the same $\hat{\beta}_2$.
- So we have two methods to solve for the same thing, what's the big deal?
  - Understanding multiple regression
  - Computational concerns
  - Understanding omitted variable bias