# 高级计量经济学
# Assignment 2

1. Show that the OLS residuals of the following regression models are numerically identical.

   (1) $\mathbf{y} = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2 + \varepsilon$,

   (2) $\mathbf{y}^* = \mathbf{X}_2^*\beta_2^* + \varepsilon^*$,

   where $\mathbf{X}_2^* = \mathbf{M}_1\mathbf{X}_2$, $\mathbf{y}^* = \mathbf{M}_1\mathbf{y}$, and $\mathbf{M}_1 = \mathbf{I} - \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'$.

2. Consider random variables $x$ and $y$ whose joint density function is $f(x, y)$. Prove the following properties.

   (1) $\mathrm{E}[xy] = \mathrm{E}_x\big[x\mathrm{E}[y\,|\,x]\big]$.

   (2) $\mathrm{Cov}[x, y] = \mathrm{Cov}_x\big[x, \mathrm{E}[y\,|\,x]\big] = \displaystyle\int_x \big(x - \mathrm{E}[x]\big)\,\mathrm{E}[y\,|\,x]\,f_x(x)\,dx.$

**Solution**

1.
We start with regression (2). The residual maker matrix of this regression is

$$\mathbf{M}^* = \mathbf{I} - \mathbf{X}_2^*(\mathbf{X}_2^{*\prime}\mathbf{X}_2^*)^{-1}\mathbf{X}_2^{*\prime}.$$

The OLS residual is then

$$\begin{aligned}
\mathbf{M}^*\mathbf{y}^* &= \mathbf{y}^* - \mathbf{X}_2^*(\mathbf{X}_2^{*\prime}\mathbf{X}_2^*)^{-1}\mathbf{X}_2^{*\prime}\mathbf{y}^* \\
&= \mathbf{M}_1\mathbf{y} - \mathbf{M}_1\mathbf{X}_2(\mathbf{X}_2'\mathbf{M}_1'\mathbf{M}_1\mathbf{X}_2)^{-1}(\mathbf{M}_1\mathbf{X}_2)'\mathbf{M}_1\mathbf{y} \\
&= \mathbf{M}_1\mathbf{y} - \mathbf{M}_1\mathbf{X}_2(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y}
\end{aligned}$$

For regression (1), from the deviation of the FWL Theorem, one sees that

$$\mathbf{b}_2 = (\mathbf{X}_2^{*\prime}\mathbf{X}_2^*)\mathbf{X}_2^{*\prime}\mathbf{y}^* = (\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y}$$

and

$$\begin{aligned}
\mathbf{b}_1 &= (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'(\mathbf{y} - \mathbf{X}_2\mathbf{b}_2) \\
&= (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\big[\mathbf{y} - \mathbf{X}_2(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y}\big] \\
&= (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\big[\mathbf{I} - \mathbf{X}_2(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_1\big]\mathbf{y}
\end{aligned}$$

Therefore, the OLS residual is

$$
\begin{aligned}
\mathbf{y}-\mathbf{X}_1\mathbf{b}_1-\mathbf{X}_2\mathbf{b}_2 &= \mathbf{y}-\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\left[\mathbf{I}-\mathbf{X}_2(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_1\right]\mathbf{y} \\
&\quad -\mathbf{X}_2(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y} \\
&= \left[\mathbf{I}-\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'+\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_1\right. \\
&\quad \left.-\mathbf{X}_2(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_1\right]\mathbf{y} \\
&= \left[\mathbf{M}_1-\mathbf{M}_1\mathbf{X}_2(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_1\right]\mathbf{y}
\end{aligned}
$$

which coincides with the OLS residual of regression (2).

2.
(1)
By definition, one has

$$
\begin{aligned}
\mathrm{E}[xy] &= \int_x\int_y xy\,f(x,y)\,dy\,dx \\
&= \int_x\int_y xy\,f(y\,|\,x)\,f_x(x)\,dy\,dx \\
&= \int_x x\left[\int_y y\,f(y\,|\,x)\,dy\right]f_x(x)\,dx \\
&= \int_x x\,\mathrm{E}[y\,|\,x]\,f_x(x)\,dx \\
&= \mathrm{E}_x\left[x\mathrm{E}[y\,|\,x]\right]
\end{aligned}
$$

(2)
The first equality follows from applying the property $\mathrm{Cov}(x,y)=\mathrm{E}[xy]-\mathrm{E}[x]\mathrm{E}[y]$ by replacing $y$ with $\mathrm{E}[y\,|\,x]$, i.e.,

$$
\begin{aligned}
\mathrm{Cov}_x(x,\mathrm{E}[y\,|\,x]) &= \mathrm{E}_x\left[x\mathrm{E}[y\,|\,x]\right]-\mathrm{E}[x]\mathrm{E}_x\left[\mathrm{E}[y\,|\,x]\right] \\
&= \mathrm{E}[xy]-\mathrm{E}[x]\mathrm{E}[y] \qquad \text{(usign (1) and Law of Iterated Expectation)} \\
&= \mathrm{Cov}[x,y]
\end{aligned}
$$

The second equality can then be derived from definition:

$$
\begin{aligned}
\mathrm{Cov}_x(x,\mathrm{E}[y\,|\,x]) &= \int_x \left(x-\mathrm{E}[x]\right)\left(\mathrm{E}[y\,|\,x]-\mathrm{E}_x\left[\mathrm{E}[y\,|\,x]\right]\right)f_x(x)\,dx \\
&= \int_x \left(x-\mathrm{E}[x]\right)\mathrm{E}[y\,|\,x]\,f_x(x)\,dx-\int_x\left(x-\mathrm{E}[x]\right)\mathrm{E}[y]\,f_x(x)\,dx \\
&= \int_x \left(x-\mathrm{E}[x]\right)\mathrm{E}[y\,|\,x]\,f_x(x)\,dx-\mathrm{E}[y]\int_x\left(x-\mathrm{E}[x]\right)f_x(x)\,dx \\
&= \int_x \left(x-\mathrm{E}[x]\right)\mathrm{E}[y\,|\,x]\,f_x(x)\,dx-\mathrm{E}[y]\cdot 0 \\
&= \int_x \left(x-\mathrm{E}[x]\right)\mathrm{E}[y\,|\,x]\,f_x(x)\,dx
\end{aligned}
$$