ECN 820B Homework 1

Willy Chen

January 25, 2023

1. Show that $Cov\left(E[X_1|X_2], X_1 - E[X_1|X_2]\right) = 0$

$$\begin{aligned}
&E[E[X_1|X_2] \cdot (X_1 - E[X_1|X_2]) \mid X_2] = E[X_1 E[X_1|X_2] - E[X_1|X_2]^2 | X_2] \\
=&E[X_1|X_2]^2 - E[X_1|X_2]^2 = 0 \\
&E[E[X_1|X_2] \cdot (X_1 - E[X_1|X_2])] = E[E[E[X_1|X_2] \cdot (X_1 - E[X_1|X_2]) \mid X_2]] \\
=&E[0|X_2] = 0 \\
&E[X_1|X_2] \cdot E[X_1 - E[X_1|X_2]] = E[X_1]E[X_1|X_2] - E[X_1|X_2] \cdot E[E[X_1|X_2]] \\
&= E[X_1]E[X_1|X_2] - E[X_1|X_2]E[X_1] = 0 \\
&Cov(E[X_1|X_2], X_1 - E[X_1|X_2]) \\
=&E[E[X_1|X_2] \cdot (X_1 - E[X_1|X_2])] - E[X_1|X_2] \cdot E[X_1 - E[X_1|X_2]] \\
=&0 - 0 = 0
\end{aligned}$$

■

2.

$$\begin{aligned}
Var(X_1 - E[X_1|X_2]) =& Cov(X_1 - E[X_1|X_2], X_1 - E[X_1|X_2]) \\
=& Cov(X_1, X_1 - E[X_1|X_2]) - \underbrace{Cov(E[X_1|X_2], X_1 - E[X_1|X_2])}_{=0 \text{ by Q1}} \\
=& Cov(X_1, X_1) - Cov(X_1, E[X_1|X_2])
\end{aligned}$$

So

$$Var(X_1 - E[X_1|X_2] \leq Var(X_1)$$

with equality if $Cov(X_1, E[X_1|X_2]) = 0$. If $X_1 \perp\!\!\!\perp X_2$, then $E[X_1|X_2] = EX_1] = c \in \mathbb{R}$, meaning that $Cov(X_1, E[X_1|X_2]) = 0$

■

3. (a) $\beta_2 = \frac{\partial E[y_t|x_t]}{\partial x_t}$

   (b) $\beta_2 \cdot \frac{1}{x_t} = E\left[\frac{1}{y_t}|x_t\right] \cdot \frac{dy_t}{dx_t} \Rightarrow \beta_2 = E[\frac{x_t}{y_t}|x_t] \cdot \frac{dy_t}{dx_t}$

(c) $\beta_2 = E\left[\frac{1}{y_t}|x_t\right] \cdot \frac{dy_t}{dx_t}$

(d) $\beta_2 \cdot \frac{1}{x_t} = \frac{\partial E[y_t|x_t]}{\partial x_t} \Rightarrow \beta_2 = x_t \cdot \frac{\partial E[y_t|x_t]}{\partial x_t}$

■

4. (a) Notice that we have

$$\begin{aligned}
married &= married * female + married * male \\
female &= single * female + married * female \\
1 &= married * female + married * male + single * female + single * male
\end{aligned}$$

(b) We can rewrite equation (1) as

$$\begin{aligned}
wage =& \beta_1 + \beta_2 educ + \beta_3 exper + \beta_4 age \\
& + \beta_5 single * female + \beta_5 married * female \\
& + \beta_6 married * female + \beta_6 married * male \\
& + \beta_7 married * female + u \\
=& \beta_1 single * male + \beta_2 educ + \beta_3 exper + \beta_4 age \\
& + (\beta_5 + \beta_1) single * female + (\beta_6 + \beta_1) married * male \\
& + (\beta_5 + \beta_6 + \beta_7 + \beta_1) married * female + u
\end{aligned}$$

So we have

$$\begin{aligned}
\beta_1 &= \delta_1 \\
\beta_5 &= \delta_2 - \delta_1 \\
\beta_6 &= \delta_3 - \delta_1 \\
\beta_7 &= \delta_4 - \delta_3 - \delta_2 + \delta_1
\end{aligned}$$

(c) We can "partial out" the effect of individual regressor on the regressand in both the standard interaction models and the fully interacted models. Additionally, so long as education, experience, and age are independent of marital status and sex, we can estimate the conditional mean model separately for the 4 "groups" and get equivalent estimates in the constants of each stratified model. This will later on lead us to specification tests such as the Hausman test.

■

5. (a) We can show this by

$$\begin{aligned}
Cov(u_t, x_t) &= E[u_t x_t] - E[u_t]E[x_t] = E[\beta_3 x_t^3 + x_t v_t] - \beta_3 E[x_t^2]E[x_t] - E[v_t]E[x_t] \\
&= \beta_3 \left(E[x_t^3] - E[x_t^2]E[x_t]\right) + E[x_t v_t] - E[x_t]E[v_t] \\
&= \beta_3 \left(\underbrace{E[x_t^3]}_{=0} - \underbrace{E[x_t^2]E[x_t]}_{=0\cdot 0}\right) + E\left[\underbrace{E[x_t v_t|x_t]}_{=x_t E[v_t|x_t]=x_t\cdot 0}\right] - E[x_t]E\left[\underbrace{E[v_t|x_t]}_{=0}\right] \\
&= 0
\end{aligned}$$

(b) We can show this by

$$E[u_t|x_t] = E[\beta_3 x_t^2 + v_t|x_t] = \beta_3 x_t^2 + E[v_t|x_t] = \beta_3 x_t^2 \neq 0$$

(c) Conditional mean independence is a stronger condition than 0 covariance (i.e. not correlated).

■

6. Notice that

$$A^{-1}ABB^{-1} = I \Rightarrow ABB^{-1}A^{-1} = AIA^{-1} = I \Rightarrow (AB)(B^{-1}A^{-1}) = I$$

Hence

$$(AB)^{-1} = B^{-1}A^{-1}$$

■

7. (a) Notice that this can be shown through an iterative process

$$(ABC\cdots)^\top = (BC\cdots)^\top A^\top = (C\cdots)^\top B^\top A^\top = \cdots C^\top B^\top A^\top$$

(b) Similarly,

$$(ABC\cdots)^{-1} = (BC\cdots)^{-1}A^{-1} = (C\cdots)^{-1}B^{-1}A^{-1} = \cdots C^{-1}B^{-1}A^{-1}$$

■

8. (a) Notice that

$$\begin{aligned} P_X^\top &= \left(X(X^\top X)^{-1}X^\top\right)^\top = X(X^{-1}(X^\top)^{-1})^\top X^\top = X(X^{-1}(X^{-1})^\top)X^\top \\ &= X(X^\top X)^{-1}X^\top \\ M_X^\top &= I^\top - P_X^\top = I - P_X = M_X \end{aligned}$$

(b) Notice that

$$\begin{aligned} P_X^2 &= X(X^\top X)^{-1}\cancel{X^\top X(X^\top X)^{-1}}X^\top = X(X^\top X)^{-1}X^\top = P_X \\ M_X^2 &= I^2 - 2P_X + P_X^2 = I - 2P_X + P_X = I - P_X = M_X \end{aligned}$$

(c) Notice that

$$\begin{aligned} P_X X &= X\cancel{(X^\top X)^{-1}X^\top X} = X \\ M_X X &= (I - P_X)X = X - P_X X = X - X = 0 \\ P_X M_X &= P_X(I - P_X) = P_X - P_X P_X = P_X - P_X = 0 \end{aligned}$$

(d) Since P_X is idempotent, we know that $tr(P_X) = rank(P_X)$, we also know that $(X^\top X)^{-1}$ is a $k \times k$ matrix and since X has full column rank, $rank\left((X^\top X)^{-1}\right) = k$, and hence $rank(P_X) = k$, so

$$\begin{aligned} tr(P_X) &= rank(P_X) = k \\ tr(M_X) &= tr(I_n - P_X) = tr(I_n) - tr(P_X) = n - k \end{aligned}$$

(e) Eigenvalue for P_X is $\lambda \in \mathbb{R}$ so $\exists v \in \mathbb{R}^n$ such that

$$P_X v = \lambda v$$

Notice that since P_X is idempotent, we have $\forall n \in \mathbb{N}$, $P_X^n = P_X$. Also, notice that given an eigenvalue $\lambda \in \mathbb{R}$ of P_X, we must have that

$$P_X^n v = \lambda^n v = P_X v = \lambda v$$

Since $\forall n \in \mathbb{N}$, $\lambda^n = \lambda$, it must be that $\lambda \in \{0, 1\}$.
Since both P_X and M_X are idempotent matrices, it must be that their eigenvalues are either 0 or 1.

■

9. (a) Notice that

$$P_{\Psi^\top X}^\top = \left(\Psi^\top X(X^\top(\Omega^{-1})^\top X)^{-1}X^\top\Psi\right)^\top \Psi^\top X(X^\top(\Omega^{-1})^\top X)^{-1}X^\top\Psi$$

Next, notice that, by construction, we have

$$(\Omega^{-1})^\top = (\Psi\Psi^\top)^\top = \Psi^{\top^\top}\Psi^\top = \Psi\Psi^\top = \Omega^{-1}$$

So we have

$$P_{\Psi^\top X}^\top = \Psi^\top X(X^\top(\Omega^{-1})^\top X)^{-1}X^\top\Psi = \Psi^\top X(X^\top\Omega^{-1}X)^{-1}$$

$$M_{\Psi^\top X}^\top = I^\top - P_{\Psi^\top X}^\top = I - P_{\Psi^\top X} = M_{\Psi^\top X}$$

So both $P_{\Psi^\top X}$ and $M_{\Psi^\top X}$ are symmetric matrices.

(b) Notice that, since $P_{\Psi^\top X}$ is symmetric, we have

$$\begin{aligned} P_{\Psi^\top X}^2 &= \Psi^\top X\cancel{(X^\top\Omega^{-1}X)^{-1}}\cancel{X^\top \underbrace{\Psi\Psi^\top}_{=\Omega^{-1}} X}(X^\top\Omega^{-1}X)^{-1}X^\top\Psi \\ &= \Psi^\top X(X^\top\Omega^{-1}X)^{-1}X^\top\Psi = P_{\Psi^\top X} \end{aligned}$$

So $P_{\Psi^\top X}$ is idempotent. Next, observe that

$$\begin{aligned} M_{\Psi^\top X}^2 &= I - 2P_{\Psi^\top X} + P_{\Psi^\top X}^2 = I - 2P_{\Psi^\top X} + P_{\Psi^\top X} \\ &= I = P_{\Psi^\top X} = M_{\Psi^\top X} \end{aligned}$$

So $M_{\Psi^\top X}$ is also idempotent.

■

10. By construction, we have that

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \sim N\left(\vec{0}, \Omega\right)$$

where $\Omega = AA^\top$ and Ω^{-1} and A^{-1} exist. Notice that this means

$$A^{-1}x \sim N\left(\vec{0}, A^{-1}\Omega(A^{-1})^\top = I_m\right)$$

We thus know that

$$x^\top(A^{-1})^\top A^{-1}x \sim \chi^2_m$$

where

$$\Omega^{-1} = (AA^\top)^{-1} = (A^\top)^{-1}A^{-1}$$

So we know that $x^\top\Omega^{-1}x \sim \chi^2_m$ ($x^\top\Omega^{-1}x$ is a chi-square random variable with m degrees of freedom.)

■

11. $z^\top P_x$ is chi-square distributed with r degrees of freedom. To see this, we must first notice that P is a symmetric and idempotent matrix with rank r (See Q8 for proof). Since P is of full rank, we can diagonalize P with $Q^\top DQ$ where $Q'Q = QQ' = I_n$ and D a diagonal matrix with r ones followed by $n-r$ zeros:

$$D = \begin{pmatrix} 1_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1_r & \ddots & 0 & 0 \\ \vdots & 0 & \ddots & 0_1 & 0 & \vdots \\ 0 & \cdots 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0_{n-r} \end{pmatrix}$$

So we can rewrite $z^\top Pz$ as

$$z^\top Pz = z^\top Q^\top DQz$$

where

$$z^\top Q^\top \sim N\left(\vec{0}, Q^\top I_n Q = Q^\top Q = I_n\right)$$

So

$$z^\top Pz = \begin{pmatrix}(zQ)_1 & \cdots & (zQ)_n\end{pmatrix} \begin{pmatrix} 1_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1_r & \ddots & 0 & 0 \\ \vdots & 0 & \ddots & 0_1 & 0 & \vdots \\ 0 & \cdots 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0_{n-r} \end{pmatrix} \begin{pmatrix}(zQ)_1 \\ \vdots \\ (zQ)_n\end{pmatrix} \sim \chi^2_r$$

■

12. By assumption, we have

$$\begin{aligned} z|X &\sim N(\vec{0}, I_n) \\ M_X &= I - X(X^\top X)^{-1}X^\top \\ M_X z &= z - X(X^\top X)^{-1}X^\top z \end{aligned}$$

So we have the covariance matrix

$$COV\left([(X^\top X)^{-1}X^\top z, M_X z] \mid X\right)$$
$$= E[(X^\top X)^{-1}X^\top zz^\top M_X^\top|X] - E[(X^\top X)^{-1}X^\top z|X]E[M_X z|X]^\top = 0_{k\times n}$$

Since

$$E[(X^\top X)^{-1}X^\top zz^\top M_X^\top|X] = (X^\top X)^{-1}X^\top E[zz^\top|X] - (X^\top X)^{-1}X^\top E[zz^\top|X]X(X^\top X)^{-1}X^\top$$
$$= (X^\top X)^{-1}X^\top (Var(z|X))M_X = (X^\top X)^{-1}X^\top I_n M_X$$
$$= (X^\top X)^{-1}X^\top - (X^\top X)^{-1}X^\top X(X^\top X)^{-1}X^\top = 0_{k\times n}$$
$$E[(X^\top X)^{-1}X^\top z|X]E[M_x z|X]^\top = (X^\top X)^{-1}X^\top \underbrace{E[z|X]}_{=\vec{0}} \underbrace{E[z|X]^\top}_{=\vec{0}} M_x$$
$$= 0_{k\times n}$$

The two vectors are conditionally uncorrelated. Since $z|X$ is normally distributed, the conditional uncorrelatedness implies that they are conditionally independent.

Similarly, we can calculate

$$COV([(X^\top X)^{-1}X^\top z, z^\top M_X z] \mid X) = \vec{0}_{n\times 1}$$
$$= E[(X^\top X)^{-1}X^\top zz^\top M_X^\top z|X] - E[(X^\top X)^{-1}X^\top z|X]E[z^\top M_X z|X]^\top$$

Since

$$E[(X^\top X)^{-1}X^\top zz^\top M_X^\top z|X] = (X^\top X)^{-1}X^\top \underbrace{E[zz^\top z|X]}_{=\vec{0}} - \underbrace{E[zz^\top X(X^\top X)^{-1}X^\top z|X]}_{=\vec{0}_{n\times 1}}$$

$$E[(X^\top X)^{-1}X^\top z|X]E[\underbrace{z^\top M_x z|X}_{\sim\chi^2_{n-k}}]^\top = (X^\top X)^{-1}X^\top \underbrace{E[z|X]}_{=\vec{0}}(n-k)$$

$$= 0_{n\times 1}$$

Once again, the two vectors are conditionally uncorrelated. Since $z|X$ is normally distributed, the conditional uncorrelatedness implies that they are conditionally independent.

The conditional independence relied on the assumption of $z|X$ being normal and the conditional covariance being 0. Generally, if $z|X$ is no longer normal, conditional independence would no longer hold. As for conditional variance, the first conditional covariance ends up being $(X^\top X)^{-1}X^\top Var(z|X)M_X$, so if $z|X$ is not distributed with a covariance matrix that is a scalar multiple of I_n, the conditional covariance may not be 0. The second conditional covariance being 0 required the fact that the odd moments of $z|X$ are 0. So if that assumption is violated, than the conditional covariance would be non-zero.

■