AS.180.633: Econometrics **Spring 2020**

Midterm: Suggested Solutions

Instructor: Yingyao Hu *By: Tong Zhou*

1. Proof. By definition,

$$\begin{aligned}\sigma^2(\boldsymbol{x}) &= \mathsf{Var}(e|\boldsymbol{x})\\ &= \mathsf{Var}\left(m(\boldsymbol{x}) + e|\boldsymbol{x}\right)\\ &= \mathsf{Var}(y|\boldsymbol{x})\\ &= \mathbb{E}[y^2|\boldsymbol{x}] - \left(\mathbb{E}[y|\boldsymbol{x}]\right)^2\end{aligned}$$

□

2. Proof. **(a)** By definition

$$\begin{aligned}Q_{\boldsymbol{xx}} &= \mathbb{E}[(\boldsymbol{xx}')]\\ &= \mathbb{E}\left[\begin{pmatrix}1\\ x_2\\ x_3\end{pmatrix}\begin{pmatrix}1 & x_2 & x_3\end{pmatrix}\right]\\ &= \begin{bmatrix}1 & \mathbb{E}(x_2) & \mathbb{E}(x_3)\\ \mathbb{E}(x_2) & \mathbb{E}(x_2^2) & \mathbb{E}(x_2x_3)\\ \mathbb{E}(x_3) & \mathbb{E}(x_2x_3) & \mathbb{E}(x_3^2)\end{bmatrix}.\end{aligned}$$

Given the condition $x_3 = \alpha_1 + \alpha_2 x_2$, the last row can be written as:

$$\begin{bmatrix}\mathbb{E}(x_3)\\ \mathbb{E}(x_2x_3)\\ \mathbb{E}(x_3^2)\end{bmatrix}^\top = \alpha_1\begin{bmatrix}1\\ \mathbb{E}(x_2)\\ \mathbb{E}(x_3)\end{bmatrix}^\top + \alpha_2\begin{bmatrix}\mathbb{E}(x_2)\\ \mathbb{E}(x_2^2)\\ \mathbb{E}(x_2x_3)\end{bmatrix}^\top$$

which suggests that the last row is a linear combination of the first two rows.

Thus, $Q_{\boldsymbol{xx}}$ is not invertible.

(b)

In what follows we use the Moore-Penrose generalized inverse, but you are not expected to do so. You can simply compute the best linear predictor of y given $(1, x_2)$.

$\boldsymbol{\beta}$, the *linear projection coefficient* of y given $\boldsymbol{x}$, satisfies $Q_{\boldsymbol{xx}}\boldsymbol{\beta} = Q_{\boldsymbol{x}y}$. Since $Q_{\boldsymbol{xx}}$ is not invertible, we resort to the Moore-Penrose generalized inverse $Q_{\boldsymbol{xx}}^-$. (See Page 831 of Hansen's lecture notes)

Let $A := \begin{pmatrix} 1 & 0 & \alpha_1 \\ 0 & 1 & \alpha_2 \end{pmatrix}$ and $\tilde{\boldsymbol{x}} := \begin{pmatrix} 1 \\ x_2 \end{pmatrix}$. Then $\boldsymbol{x} = A'\tilde{\boldsymbol{x}}$ and $Q_{\boldsymbol{xx}} = \mathbb{E}(\boldsymbol{xx}') = A'\mathbb{E}(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}')A$.

In what follows we assume x_2 is nondegenerate (i.e. $\mathsf{Var}(x_2) > 0$), which implies $\mathbb{E}(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}')$ is invertible, and the Moore-Penrose generalized inverse

$$\begin{aligned} Q^{-}_{\boldsymbol{xx}} &= A^{-}\left(\mathbb{E}(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}')\right)^{-}\left(A'\right)^{-} \\ &= A^{-}\left(\mathbb{E}(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}')\right)^{-1}\left(A'\right)^{-}. \end{aligned}$$

Where,

$$\begin{aligned} A^{-} = A'\left(AA'\right)^{-1} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \alpha_1 & \alpha_2 \end{bmatrix}\begin{bmatrix} 1+\alpha_1^2 & \alpha_1\alpha_2 \\ \alpha_1\alpha_2 & 1+\alpha_2^2 \end{bmatrix}^{-1} \\ \left(A'\right)^{-} = \left(AA'\right)^{-1}A &= \begin{bmatrix} 1+\alpha_1^2 & \alpha_1\alpha_2 \\ \alpha_1\alpha_2 & 1+\alpha_2^2 \end{bmatrix}^{-1}\begin{bmatrix} 1 & 0 & \alpha_1 \\ 0 & 1 & \alpha_2 \end{bmatrix} \\ \left(\mathbb{E}(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}')\right)^{-1} &= \begin{bmatrix} 1 & \mathbb{E}(x_2) \\ \mathbb{E}(x_2) & \mathbb{E}(x_2^2) \end{bmatrix}^{-1}. \end{aligned}$$

Note that you should convince yourself that A^{-} and $\left(A'\right)^{-}$ are indeed the Moore-Penrose generalized inverses of A and A' by checking four conditions listed in Hansen's lecture notes.

Then $\boldsymbol{\beta} = Q^{-}_{\boldsymbol{xx}}Q_{\boldsymbol{x}y}$, and the best linear predictor:

$$\begin{aligned} \mathcal{P}(y|\boldsymbol{x}) &= \boldsymbol{x}'\boldsymbol{\beta} \\ &= \boldsymbol{x}'Q^{-}_{\boldsymbol{xx}}Q_{\boldsymbol{x}y} \\ &= \frac{\mathbb{E}(y)\mathbb{E}(x_2^2) - \mathbb{E}(x_2)\mathbb{E}(x_2y) + x_2\mathsf{Cov}(x_2,y)}{\mathsf{Var}(x_2)} \\ &= \mathbb{E}(y) + (x_2 - \mathbb{E}(x_2))\underbrace{\frac{\mathsf{Cov}(x_2,y)}{\mathsf{Var}(x_2)}}_{\text{define } \beta^* :=} \\ &= \mathbb{E}(y) + (x_2 - \mathbb{E}(x_2))\beta^* \\ &= \underbrace{\mathbb{E}(y) - \mathbb{E}(x_2)\beta^*}_{\text{define } \alpha^* :=} + x_2\beta^* \\ &= \alpha^* + x_2\beta^* \\ &= \tilde{\boldsymbol{x}}'\tilde{\boldsymbol{\beta}} \end{aligned}$$

where

$$
\begin{aligned}
Q_{xy} &= \mathbb{E}(\boldsymbol{x}y) = A'\mathbb{E}(\tilde{\boldsymbol{x}}y) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \alpha_1 & \alpha_2 \end{bmatrix} \begin{bmatrix} 1 \\ \mathbb{E}(x_2) \end{bmatrix} \\
\beta^* &:= \frac{\mathsf{Cov}(x_2, y)}{\mathsf{Var}(x_2)} \\
\alpha^* &:= \mathbb{E}(y) - \beta^*\mathbb{E}(x_2) \\
\tilde{\boldsymbol{\beta}} &:= \begin{bmatrix} \alpha^* & \beta^* \end{bmatrix}^\top .
\end{aligned}
$$

Moreover, It is straightforward that $\tilde{\boldsymbol{\beta}}$ is the *linear projection coefficient* of y given $\tilde{\boldsymbol{x}}$ (i.e. $\mathcal{P}(y|\tilde{\boldsymbol{x}}) = \tilde{\boldsymbol{x}}'\tilde{\boldsymbol{\beta}}$). Therefore, $\mathcal{P}(y|\boldsymbol{x})$ agrees with $\mathcal{P}(y|\tilde{\boldsymbol{x}})$:

$$
\mathcal{P}(y|\boldsymbol{x}) = \mathcal{P}(y|\tilde{\boldsymbol{x}}) = \tilde{\boldsymbol{x}}'\tilde{\boldsymbol{\beta}} = \alpha^* + x_2\beta^*.
$$

where α^* and β^* defined as above. □

3. Proof. By

$$
\begin{aligned}
\boldsymbol{P} &= \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \\
&= \begin{bmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{X}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{X}_2'\boldsymbol{X}_2 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{X}_1' \\ \boldsymbol{X}_2' \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{bmatrix} \begin{bmatrix} \left(\boldsymbol{X}_1'\boldsymbol{X}_1\right)^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \left(\boldsymbol{X}_2'\boldsymbol{X}_2\right)^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_1' \\ \boldsymbol{X}_2' \end{bmatrix} \\
&= \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' + \boldsymbol{X}_2(\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2' \\
&= \boldsymbol{P}_1 + \boldsymbol{P}_2
\end{aligned}
$$

Thus, $\boldsymbol{P} = \boldsymbol{P}_1 + \boldsymbol{P}_2$. □

4. Proof. By (3.44) of Hansen's notes,

$$
\widehat{\boldsymbol{\beta}}_{(-i)} = \widehat{\boldsymbol{\beta}} - \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{x}_i(1 - h_{ii})^{-1}\widehat{e}_i
$$

So the condition under which $\widehat{\boldsymbol{\beta}}_{(-i)} = \widehat{\boldsymbol{\beta}}$ is that $\widehat{e}_i = 0$ for some i.

□

5. Proof. **(a)** We know

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \\ &= \left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{y}\right) \\ &= \frac{1}{n}\boldsymbol{X}'\boldsymbol{y} \\ &= \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i y_i.\end{aligned}$$

So,

$$\begin{aligned}\boldsymbol{V}_{\hat{\boldsymbol{\beta}}} &= \frac{1}{n^2}\operatorname{Var}\left(\sum_{i=1}^{n}\boldsymbol{x}_i y_i\right) \\ &= \frac{1}{n^2}\left(\sum_{i=1}^{n}\operatorname{Var}(\boldsymbol{x}_i y_i) + \sum_{i\neq j}\underbrace{\operatorname{Cov}\left(\boldsymbol{x}_i y_i, \boldsymbol{x}_j y_j\right)}_{=0}\right) \\ &= \frac{1}{n^2}\sum_{i=1}^{n}\operatorname{Var}(\boldsymbol{x}_i y_i) \qquad (1)\end{aligned}$$

However, for the condition

$$\frac{\boldsymbol{X}'\boldsymbol{X}}{n} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i' = \boldsymbol{I}_k,$$

taking expectation and variance on both sides gives:

$$\begin{aligned}\mathbb{E}[\boldsymbol{x}_i\boldsymbol{x}_i'] &= \boldsymbol{I}_k \\ \operatorname{Var}\left(\boldsymbol{x}_i\boldsymbol{x}_i'\right) &= \boldsymbol{0}.\end{aligned}$$

Then for any single term $\operatorname{Var}(\boldsymbol{x}_i y_i)$ in Eq. (1), by the law of total variance

$$\begin{aligned}\operatorname{Var}(\boldsymbol{x}_i y_i) &= \mathbb{E}[\operatorname{Var}(\boldsymbol{x}_i y_i|\boldsymbol{x}_i)] + \operatorname{Var}\left(\mathbb{E}[\boldsymbol{x}_i y_i|\boldsymbol{x}_i]\right) \\ &= \mathbb{E}[\operatorname{Var}(\boldsymbol{x}_i e_i|\boldsymbol{x}_i)] + \boldsymbol{\beta}\underbrace{\operatorname{Var}(\boldsymbol{x}_i\boldsymbol{x}_i')}_{=\boldsymbol{0}}\boldsymbol{\beta}' \\ &= \sigma_i^2\boldsymbol{I}_k\end{aligned}$$

Plugging it into Eq. (1), we obtain

$$\boldsymbol{V}_{\hat{\boldsymbol{\beta}}} = \left(\frac{\sum_{i=1}^{n}\sigma_i^2}{n^2}\right)\boldsymbol{I}_k.$$

Therefore, $\boldsymbol{V}_{\hat{\boldsymbol{\beta}}}$ is diagonal.

(b) Diagonal $\boldsymbol{V}_{\hat{\boldsymbol{\beta}}}$ implies that $\hat{\beta}_j$ and $\hat{\beta}_\ell$ are uncorrelated.

(c) Even if we do not make further assumptions (e.g. homoskedasticity), $\widehat{\beta}_j$ and $\widehat{\beta}_\ell$ are automatically uncorrelated for any $j \neq \ell$. □

6. Proof. **(a)** By definition, $\mathsf{Cov}\left(\widetilde{\boldsymbol{\beta}}, \widehat{\boldsymbol{\beta}} \middle| \boldsymbol{X}\right) = \mathbb{E}[(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta})' | \boldsymbol{X}]$, then copy the proof of Problem 4.20 of Homework 4 and therefore

$$\mathsf{Cov}\left(\widehat{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\beta}} \middle| \boldsymbol{X}\right) = \mathsf{Var}\left(\widetilde{\boldsymbol{\beta}} \mid \boldsymbol{X}\right)$$

and

$$\mathsf{Cov}\left(\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}} \middle| \boldsymbol{X}\right) = \mathsf{Var}\left(\widetilde{\boldsymbol{\beta}} \middle| \boldsymbol{X}\right) - \mathsf{Cov}\left(\widetilde{\boldsymbol{\beta}}, \widehat{\boldsymbol{\beta}} \middle| \boldsymbol{X}\right) = \mathbf{0}.$$

(b) It can be proved by the following orthogonal matrices:

$$\begin{aligned}
\boldsymbol{M} &= \boldsymbol{I} - \boldsymbol{P} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \\
\boldsymbol{M}_1 &= \boldsymbol{I} - \boldsymbol{P}_1 = \boldsymbol{I} - \boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}
\end{aligned}$$

Then

$$\begin{aligned}
\hat{\boldsymbol{e}} &= \boldsymbol{M}\boldsymbol{e} \\
\tilde{\boldsymbol{e}} &= \boldsymbol{M}_1\boldsymbol{e}
\end{aligned}$$

It is easy to verify that $\boldsymbol{M}_1'\boldsymbol{M}\boldsymbol{M}_1 = \boldsymbol{M}$. Therefore,

$$\begin{aligned}
\tilde{\boldsymbol{e}}'\tilde{\boldsymbol{e}} - \hat{\boldsymbol{e}}'\hat{\boldsymbol{e}} &= \boldsymbol{e}'\boldsymbol{M}_1'\boldsymbol{M}_1\boldsymbol{e} - \boldsymbol{e}'\boldsymbol{M}\boldsymbol{e} \\
&= \boldsymbol{e}'\left(\boldsymbol{M}_1'\boldsymbol{M}_1 - \boldsymbol{M}\right)\boldsymbol{e} \\
&= \boldsymbol{e}'\left(\boldsymbol{M}_1'\boldsymbol{M}_1 - \boldsymbol{M}_1'\boldsymbol{M}\boldsymbol{M}_1\right)\boldsymbol{e} \\
&= (\boldsymbol{M}_1\boldsymbol{e})'(\boldsymbol{I} - \boldsymbol{M})(\boldsymbol{M}_1\boldsymbol{e}) \\
&= (\boldsymbol{M}_1\boldsymbol{e})'\boldsymbol{P}(\boldsymbol{M}_1\boldsymbol{e}) \\
&= (\boldsymbol{P}\boldsymbol{M}_1\boldsymbol{e})'(\boldsymbol{P}\boldsymbol{M}_1\boldsymbol{e}) \\
&\geqslant 0,
\end{aligned}$$

thus, $\hat{R}^2 \geqslant \tilde{R}^2$. That is, even if the heteroskedasticity is present, $\hat{R}^2$ is still greater.

The intuition is that $\boldsymbol{M}_1$ is idempotent but not symmetric, so that $\tilde{\boldsymbol{y}}$ is an *oblique* (but not orthogonal) projection of $\boldsymbol{y}$ onto span($\boldsymbol{X}$). Then it follows that $\hat{\boldsymbol{e}}'\hat{\boldsymbol{e}} \leqslant \tilde{\boldsymbol{e}}'\tilde{\boldsymbol{e}}$. □

7. Proof. See Problem 5.5 of Homework 5. □

8. Proof. **(a)** Since

$$\frac{(n-k)s^2}{\sigma^2} \sim \chi^2_{n-k},$$

$\mathrm{Var}\left(\frac{(n-k)s^2}{\sigma^2}\right) = 2(n-k)$, which implies $\mathrm{Var}(s^2) = \frac{2\sigma^4}{(n-k)}$.

The Cramér-Rao lower bound for σ^2 is $2\sigma^4/n$ that is strictly smaller than $\mathrm{Var}(s^2)$.

(b) When the homoskedasticity is true (i.e. $\mathrm{Var}(\boldsymbol{e}|\boldsymbol{X}) = \sigma^2\boldsymbol{I}_n$), $\mathrm{Var}\left(\widehat{\boldsymbol{\beta}}\,\middle|\,\boldsymbol{X}\right) = \sigma^2\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}$ which is the Cramér-Rao lower bound for $\boldsymbol{\beta}$. □