Additional Practics Questions for Midterm 1 Solutions

Hansen 3.2

Let's call $\tilde{\beta}$ and $\tilde{\mathbf{e}}$ the OLS estimates and residuals from the regression of $\mathbf{Y}$ on $\mathbf{Z}$. Notice that

$$\begin{aligned}
\tilde{\beta} &= (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y} \\
&= \left((\mathbf{XC})'\mathbf{XC}\right)^{-1}(\mathbf{XC})'\mathbf{Y} \\
&= (\mathbf{C}'\mathbf{X}'\mathbf{XC})^{-1}\mathbf{C}'\mathbf{X}'\mathbf{Y} \\
&= \mathbf{C}^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'^{-1}\mathbf{C}'\mathbf{X}'\mathbf{Y} \\
&= \mathbf{C}^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \\
&= \mathbf{C}^{-1}\hat{\beta}
\end{aligned}$$

where the second equality holds by plugging in $\mathbf{Z} = \mathbf{XC}$, the third equality holds by taking the transpose of $\mathbf{XC}$, the fourth equality holds because $\mathbf{C}$ and $\mathbf{X}'\mathbf{X}$ are nonsingular, the fifth equality holds by canceling the $\mathbf{C}'^{-1}\mathbf{C}'$, and the last equality holds by the definition of $\hat{\beta}$.

Now, for the residuals, notice that

$$\begin{aligned}
\tilde{\mathbf{e}} &= \mathbf{Y} - \mathbf{Z}\tilde{\beta} \\
&= \mathbf{Y} - \mathbf{XCC}^{-1}\hat{\beta} \\
&= \mathbf{Y} - \mathbf{X}\hat{\beta} \\
&= \hat{\mathbf{e}}
\end{aligned}$$

where this result holds just by plugging in and canceling terms. This says that the residuals from the regression of $\mathbf{Y}$ on $\mathbf{Z}$ are exactly the same as the residuals from the regression of $\mathbf{Y}$ on $\mathbf{X}$.

As a side-comment, a simple example of this problem would be something like scaling all the regressors by, say, 100. If you did this, it would change the value of the estimated coefficients (divide them by 100) but would fit the data equally well.

Hansen 3.5

The OLS coefficient from a regression of $\hat{\mathbf{e}}$ on $\mathbf{X}$ is given by

$$\begin{aligned}
(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{e}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{Y} - \mathbf{X}\hat{\beta}) \\
&= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\hat{\beta} \\
&= \hat{\beta} - \hat{\beta} \\
&= 0
\end{aligned}$$

where the first part of the third equality holds by the definition of $\hat{\beta}$ and the last part holds by canceling the terms involving $(\mathbf{X}'\mathbf{X})$.

Hansen 3.6

Call $\hat{\gamma}$ the coefficient of the regression of $\hat{\mathbf{Y}}$ on $\mathbf{X}$. It is given by

$$\begin{aligned}\hat{\gamma} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{Y}} \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\hat{\beta} \\ &= \hat{\beta}\end{aligned}$$

so that running a regression of $\hat{\mathbf{Y}}$ on $\mathbf{X}$ gives the same coefficient as running a regression of $\mathbf{Y}$ on $\mathbf{X}$.

Hansen 3.7

$$\begin{aligned}\mathbf{P}\mathbf{X} &= \mathbf{P}\begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{P}\mathbf{X}_1 & \mathbf{P}\mathbf{X}_2\end{bmatrix}\end{aligned}$$

Further, since $\mathbf{X} = \begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2\end{bmatrix}$ and $\mathbf{P}\mathbf{X} = \mathbf{X}$ (from the properties of the projection matrix $\mathbf{P}$), this implies that $\mathbf{P}\mathbf{X}_1 = \mathbf{X}_1$.

Similarly,

$$\begin{aligned}\mathbf{M}\mathbf{X} &= \mathbf{M}\begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{M}\mathbf{X}_1 & \mathbf{M}\mathbf{X}_2\end{bmatrix}\end{aligned}$$

but we also know that $\mathbf{M}\mathbf{X} = \mathbf{0}_{n\times k} = \begin{bmatrix}\mathbf{0}_{n\times k_1} & \mathbf{0}_{n\times k_2}\end{bmatrix}$ where, for example, $\mathbf{0}_{n\times k_1}$ is an $n \times k_1$ matrix of zeroes. This implies that $\mathbf{M}\mathbf{X}_1 = \mathbf{0}_{n\times k_1}$.

Hansen 3.10

Notice that

$$\begin{aligned}\mathbf{P} &= \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \\ &= \begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2\end{bmatrix}\left(\begin{bmatrix}\mathbf{X}_1' \\ \mathbf{X}_2'\end{bmatrix}\begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2\end{bmatrix}\right)^{-1}\begin{bmatrix}\mathbf{X}_1' \\ \mathbf{X}_2'\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2\end{bmatrix}\left(\begin{bmatrix}\mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\mathbf{X}_2\end{bmatrix}\right)^{-1}\begin{bmatrix}\mathbf{X}_1' \\ \mathbf{X}_2'\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2\end{bmatrix}\left(\begin{bmatrix}\mathbf{X}_1'\mathbf{X}_1 & \mathbf{0}_{k_1\times k_2} \\ \mathbf{0}_{k_2\times k_1} & \mathbf{X}_2'\mathbf{X}_2\end{bmatrix}\right)^{-1}\begin{bmatrix}\mathbf{X}_1' \\ \mathbf{X}_2'\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}(\mathbf{X}_1'\mathbf{X}_1)^{-1} & \mathbf{0}_{k_1\times k_2} \\ \mathbf{0}_{k_2\times k_1} & (\mathbf{X}_2'\mathbf{X}_2)^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{X}_1' \\ \mathbf{X}_2'\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1} & \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{X}_1' \\ \mathbf{X}_2'\end{bmatrix} \\ &= \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1' + \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2' \\ &= \mathbf{P}_1 + \mathbf{P}_2\end{aligned}$$

where the first equality holds by the definition of $\mathbf{P}$, the second equality holds by partitioning $\mathbf{X}$ as in the problem, the third equality holds by multiplying the two matrices inside the inverse, the fourth equality holds because $\mathbf{X}_1'\mathbf{X}_2 = 0$, the fifth equality holds because the inverse of a block diagonal matrix is equal to the inverse of the blocks, the sixth equality holds by multiplying the first two matrices, the seventh equality holds again by matrix multiplication, and the last equality holds by the definition of $\mathbf{P}_1$ and $\mathbf{P}_2$.

Hansen 3.22

From the first regression, we immediately have that

$$\tilde{\mathbf{u}} = \mathbf{M}_1\mathbf{Y}$$

which holds because it is a regression of Y on X_1 (and I use bold font above to indicate that, e.g., $\tilde{\mathbf{u}}$ is the $n \times 1$ vector of residuals from the first regression). Then, the coefficient from the second regression is given by

$$\begin{aligned}\tilde{\beta}_2 &= (\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'\tilde{\mathbf{u}} \\ &= (\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{Y}\end{aligned}$$

We can compare this to $\hat{\beta}$ from the third regression given in the problem. We immediately know from FWL-type arguments that

$$\hat{\beta}_2 = (\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{Y}$$

In general, these are not equal to each other.

Hansen 4.1

Part a

$$\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} Y_i^k$$

Part b

$$\begin{aligned}\mathbb{E}\left[\hat{\mu}_k\right] &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} Y_i^k\right] \\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[Y_i^k] \\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[Y^k] \\ &= \mathbb{E}[Y^k]\end{aligned}$$

where the third equality holds because the Y_i are identically distributed (implying the mean is the same across i). This result implies that $\hat{\mu}_k$ is unbiased for μ_k.

Part c

$$
\begin{aligned}
\text{var}\left(\hat{\mu}_k\right) &= \text{var}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i^k\right) \\
&= \frac{1}{n^2}\text{var}\left(\sum_{i=1}^{n} Y_i^k\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\text{var}(Y^k) \\
&= \frac{\text{var}(Y^k)}{n}
\end{aligned}
$$

where the second equality holds because $1/n$ is a constant and it should be squared to come out of the variance, the third equality holds by passing the variance through the sum (in order for their not to be any covariance terms introduced here, it requires the "independence" part of iid; for this variance to be the same across all units requires the "identically distributed" part of iid), and the last equality holds because summing a constant n times cancels one of the n's from the denominator.

For $\text{var}(\hat{\mu}_k)$ to exist, we need for $\text{var}(Y^k)$ to exist. Notice that,

$$
\text{var}(Y^k) = \mathbb{E}[(Y^k)^2] - \mathbb{E}[Y^k]^2
$$

Thus, the condition that we need is that $\mathbb{E}[(Y^k)^2] = \mathbb{E}[Y^{2k}] < \infty$.

Part d

We can estimate by

$$
\widehat{\text{var}}(\hat{\mu}_k) = \frac{\widehat{\text{var}}(Y^k)}{n} = \frac{\frac{1}{n}\sum_{i=1}^{n} Y_i^{2k} - \left(\frac{1}{n}\sum_{i=1}^{n} Y_i^k\right)^2}{n}
$$

Hansen 4.5

First (and notice that this is exactly the same as what we showed in class...because unbiasedness did not rely on homoskedasticity),

$$
\begin{aligned}
\mathbb{E}[\hat{\beta}|\mathbf{X}] &= \mathbb{E}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}|\mathbf{X}] \\
&= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbb{E}[\mathbf{Y}|\mathbf{X}] \\
&= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\beta \\
&= \beta
\end{aligned}
$$

For thinking about the sampling variance, first notice that

$$\begin{aligned}\mathrm{var}(\mathbf{Y}|\mathbf{X}) &= \mathrm{var}(\mathbf{X}\beta + \mathbf{e}|\mathbf{X}) \\ &= \mathrm{var}(\mathbf{e}|\mathbf{X}) \\ &= \sigma^2\mathbf{\Sigma}\end{aligned}$$

where the first equality holds by plugging in for $\mathbf{Y}$, the second equality holds because we are conditioning on $\mathbf{X}$, and the third equality holds from the way that $\mathbf{\Sigma}$ is defined in the textbook. Next,

$$\begin{aligned}\mathrm{var}(\hat{\beta}|\mathbf{X}) &= \mathrm{var}\Big((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}|\mathbf{X}\Big) \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathrm{var}(\mathbf{Y}|\mathbf{X})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\end{aligned}$$

where the first equality holds by plugging in for $\hat{\beta}$, the second equality holds because the variance is conditional on $\mathbf{X}$ (so the terms involving $\mathbf{X}$ can come out but need to be "squared"), and the last equality holds by plugging in the expression for $\mathrm{var}(\mathbf{Y}|\mathbf{X})$ that we derived above. This is the result we were trying to show.

Hansen 4.23

Notice that

$$\begin{aligned}\mathbb{E}[\hat{\beta}_{ridge}|\mathbf{X}] &= \mathbb{E}\left[(\mathbf{X}'\mathbf{X} + \mathbf{I}_k\lambda)^{-1}\mathbf{X}'\mathbf{Y}\right] \\ &= (\mathbf{X}'\mathbf{X} + \mathbf{I}_k\lambda)^{-1}\mathbf{X}'\mathbb{E}[\mathbf{Y}|\mathbf{X}] \\ &= (\mathbf{X}'\mathbf{X} + \mathbf{I}_k\lambda)^{-1}\mathbf{X}'\mathbf{X}\beta \\ &\neq \beta\end{aligned}$$

This implies that $\hat{\beta}_{ridge}$ is not unbiased for β.