Homework 5

The due date is December 12, Wednesday.

Problem 1. Consider the linear regression model

$$\begin{aligned} \boldsymbol{Y} &= \boldsymbol{X\beta} + \boldsymbol{e}, \\ \mathbb{E}(\boldsymbol{e}|\boldsymbol{X}) &= 0, \\ \operatorname{rank}(\boldsymbol{X}) &= k, \\ \mathbb{E}(\boldsymbol{ee}'|\boldsymbol{X}) &= \sigma^2 \boldsymbol{I}_n, \end{aligned}$$

where $\boldsymbol{X}$ is the $n \times k$ matrix of regressors, $\boldsymbol{Y}$ is the n-vector of observations on the dependent variable, and $\boldsymbol{\beta} \in \mathbb{R}^k$ and $\sigma^2 > 0$ are unknown parameters. In addition, assume that β *satisfies the restriction*

$$\boldsymbol{R\beta} - \boldsymbol{r} = \boldsymbol{0},$$

where $\boldsymbol{R}$ is a known non-random $q \times k$ matrix of rank q, and $\boldsymbol{r}$ is a known non-random q-vector. Consider the restricted LS estimator

$$\widetilde{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}} - \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{R}' \left(\boldsymbol{R}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{R}'\right)^{-1} \left(\boldsymbol{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right),$$

where $\widehat{\boldsymbol{\beta}}$ is the unrestricted LS estimator

$$\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{X}'\boldsymbol{Y}.$$

(i) Show that $\widetilde{\boldsymbol{\beta}}$ can be written as

$$\widetilde{\boldsymbol{\beta}} = \boldsymbol{\beta} + (\boldsymbol{I}_n - \boldsymbol{QR})\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{X}'\boldsymbol{e},$$

where

$$\boldsymbol{Q} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{R}' \left(\boldsymbol{R}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{R}'\right)^{-1}.$$

(ii) Is $\widetilde{\boldsymbol{\beta}}$ unbiased?

(iii) Show that $\operatorname{Var}\left(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}\right) = \sigma^2 (\boldsymbol{I}_n - \boldsymbol{QR})\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} (\boldsymbol{I}_n - \boldsymbol{QR})'$.

(iv) Show that $\boldsymbol{QR}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{R}'\boldsymbol{Q}' = \boldsymbol{QR}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{R}'\boldsymbol{Q}'$.

(v) Using the results from parts (iii) and (iv), show that $\operatorname{Var}\left(\widehat{\boldsymbol{\beta}}|\boldsymbol{X}\right) - \operatorname{Var}\left(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}\right) \geq 0$ (in the positive semidefinite sense).

Solution. (i)

$$\begin{aligned} \widetilde{\boldsymbol{\beta}} &= \widehat{\boldsymbol{\beta}} - \boldsymbol{Q}(\boldsymbol{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}) \\ &= \widehat{\boldsymbol{\beta}} - \boldsymbol{QR}\widehat{\boldsymbol{\beta}} + \boldsymbol{Qr} \\ &= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{X}'(\boldsymbol{X\beta} + \boldsymbol{e}) - \boldsymbol{QR}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{X}'(\boldsymbol{X\beta} + \boldsymbol{e}) + \boldsymbol{Qr} \\ &= \boldsymbol{\beta} + \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{X}'\boldsymbol{e} - \boldsymbol{QR\beta} - \boldsymbol{QR}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{X}'\boldsymbol{e} + \boldsymbol{Qr} \\ &= \boldsymbol{\beta} + (\boldsymbol{I}_n - \boldsymbol{QR})\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{X}'\boldsymbol{e} - \boldsymbol{Q}(\boldsymbol{R\beta} - \boldsymbol{r}) \end{aligned}$$

$$= \boldsymbol{\beta} + (\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{e}.$$

(ii)

$$\begin{aligned}
\mathbb{E}\left(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}\right) &= \mathbb{E}\left(\boldsymbol{\beta} + (\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{e}|\boldsymbol{X}\right) \\
&= \boldsymbol{\beta} + (\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\mathbb{E}\left(\boldsymbol{e}|\boldsymbol{X}\right) \\
&= \boldsymbol{\beta}.
\end{aligned}$$

By LIE $\mathbb{E}(\widetilde{\boldsymbol{\beta}}) = \mathbb{E}\mathbb{E}\left(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}\right) = \boldsymbol{\beta}$.

(iii)

$$\begin{aligned}
\operatorname{Var}\left(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}\right) &= \operatorname{Var}\left(\boldsymbol{\beta} + (\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{e}|\boldsymbol{X}\right) \\
&= (\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\operatorname{Var}\left(\boldsymbol{e}|\boldsymbol{X}\right)\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}(\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})' \\
&= (\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\mathbb{E}\left(\boldsymbol{e}\boldsymbol{e}'|\boldsymbol{X}\right)\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}(\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})' \\
&= (\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\sigma^2\boldsymbol{I}_n\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}(\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})' \\
&= \sigma^2(\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}(\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})'.
\end{aligned}$$

(iv)

$$\begin{aligned}
\boldsymbol{Q} &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'(\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}')^{-1} \\
\boldsymbol{Q}' &= (\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}')^{-1}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}
\end{aligned}$$

Therefore

$$\boldsymbol{Q}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\underbrace{(\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}')^{-1}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}}_{\boldsymbol{Q}'} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{Q}'$$

$$\begin{aligned}
\boldsymbol{Q}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{Q}' &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'(\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}')^{-1}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'(\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}')^{-1}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\underbrace{(\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}')^{-1}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}}_{\boldsymbol{Q}'} \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{Q}' = \boldsymbol{Q}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}
\end{aligned}$$

(v)

$$\begin{aligned}
&\operatorname{Var}\left(\widehat{\boldsymbol{\beta}}|\boldsymbol{X}\right) - \operatorname{Var}\left(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}\right) \\
=&\sigma^2\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} - \sigma^2(\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}(\boldsymbol{I}_n - \boldsymbol{Q}\boldsymbol{R})' \\
=&\sigma^2\left[(\boldsymbol{X}'\boldsymbol{X})^{-1} - (\boldsymbol{X}'\boldsymbol{X})^{-1} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{Q}' + \boldsymbol{Q}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1} - \boldsymbol{Q}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{Q}'\right] \\
=&\sigma^2\left[\boldsymbol{Q}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{Q}' + \boldsymbol{Q}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{Q}' - \boldsymbol{Q}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{Q}'\right] \\
=&\sigma^2\boldsymbol{Q}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{Q}'
\end{aligned}$$

Since $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ is positive definite, for any $\boldsymbol{Q}\boldsymbol{R}$, $\boldsymbol{Q}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{Q}\boldsymbol{R})' \geq 0$. Therefore we have $\operatorname{Var}\left(\widehat{\boldsymbol{\beta}}|\boldsymbol{X}\right) - \operatorname{Var}\left(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}\right) \geq 0$.

Problem 2. Suppose we have a simple regression model:

$$Y_i = \beta_0 + \beta_1 X_i + U_i.$$

Assume that the observations are iid but $\mathbb{E}\left(X_i U_i\right) \neq 0$.

(i) Show that the LS estimator $\widehat{\beta}_1$ is not a consistent estimator for β_1 so $p\lim \hat{\beta}_1 \neq \beta_1$.

(ii) Suppose we observe a variable Z_i such that $\mathbb{E}(U_i|Z_i) = 0$ and $\operatorname{Cov}(X_i, Z_i) \neq 0$. Show that

$$\widetilde{\beta}_1 = \frac{\sum_{i=1}^n \left(Z_i - \overline{Z}\right) Y_i}{\sum_{i=1}^n \left(Z_i - \overline{Z}\right) X_i}$$

is a consistent estimator for β_1. $\overline{Z} = n^{-1} \sum_{i=1}^n Z_i$.

Solution. The LS estimator:

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n \left(X_i - \overline{X}\right) Y_i}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} = \frac{\sum_{i=1}^n \left(X_i - \overline{X}\right)(\beta_0 + \beta_1 X_i + U_i)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} = \beta_1 + \frac{n^{-1}\sum_{i=1}^n \left(X_i - \overline{X}\right) U_i}{n^{-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}.$$

By WLLN and Continuous Mapping Theorem,

$$n^{-1}\sum_{i=1}^n \left(X_i - \overline{X}\right) U_i = n^{-1}\sum_{i=1}^n X_i U_i - \overline{X} n^{-1} \sum_{i=1}^n U_i \to_p \mathbb{E}(X_i U_i) - \mathbb{E}(X_i)\mathbb{E}(U_i) \neq 0$$

and

$$n^{-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2 = n^{-1}\sum_{i=1}^n X_i^2 - \overline{X}^2 \to_p \mathbb{E}\left(X_i^2\right) - \mathbb{E}(X_i)^2 = \operatorname{Var}(X_i).$$

So,

$$\frac{n^{-1}\sum_{i=1}^n \left(X_i - \overline{X}\right) U_i}{n^{-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} \to_p \frac{\mathbb{E}(X_i U_i) - \mathbb{E}(X_i)\mathbb{E}(U_i)}{\operatorname{Var}(X_i)} \neq 0.$$

And

$$\widetilde{\beta}_1 = \frac{\sum_{i=1}^n \left(Z_i - \overline{Z}\right) Y_i}{\sum_{i=1}^n \left(Z_i - \overline{Z}\right) X_i} = \frac{\sum_{i=1}^n \left(Z_i - \overline{Z}\right)(\beta_0 + \beta_1 X_i + U_i)}{\sum_{i=1}^n \left(Z_i - \overline{Z}\right) X_i} = \beta_1 + \frac{n^{-1}\sum_{i=1}^n \left(Z_i - \overline{Z}\right) U_i}{n^{-1}\sum_{i=1}^n \left(Z_i - \overline{Z}\right)\left(X_i - \overline{X}\right)}.$$

Similarly,

$$\frac{n^{-1}\sum_{i=1}^n \left(Z_i - \overline{Z}\right) U_i}{n^{-1}\sum_{i=1}^n \left(Z_i - \overline{Z}\right)\left(X_i - \overline{X}\right)} \to_p \frac{\mathbb{E}(Z_i U_i) - \mathbb{E}(Z_i)\mathbb{E}(U_i)}{\operatorname{Cov}(X_i, Z_i)} = 0.$$

Problem 3. Aggregate demand Q_D for a certain commodity is determined by its price P, aggregate income Y, and population, POP,

$$Q_D = \beta_1 + \beta_2 P + \beta_3 Y + \beta_4 POP + U^D$$

and aggregate supply is given by

$$Q_S = \alpha_1 + \alpha_2 P + U^S$$

where U_D and U_S are independently distributed error terms: U_D and U_S are independent from all other variables and they are also independent from each other. Remember that the quantity and the price are determined simultaneously in the equilibrium $Q_S = Q_D = Q$. We observe only the equilibrium values Q so that the observed price must satisfy the equation (demand = supply):

$$\beta_1 + \beta_2 P + \beta_3 Y + \beta_4 POP + U^D = \alpha_1 + \alpha_2 P + U^S.$$

(i) Show that the LS estimator of α_2 will be inconsistent if LS is used to fit the supply equation.

(ii) Show that a consistent estimator of α_2 is

$$\widetilde{\alpha}_2 = \frac{\sum_{i=1}^n \left(Y_i - \overline{Y}\right)\left(Q_i - \overline{Q}\right)}{\sum_{i=1}^n \left(Y_i - \overline{Y}\right)\left(P_i - \overline{P}\right)}.$$

Solution. The reduced form equation (which expresses P as a function of the explanatory variables and the error terms) for P is

$$P = \frac{1}{\alpha_2 - \beta_2}\left(\beta_1 - \alpha_1 + \beta_3 Y + \beta_4 POP + U^D - U^S\right).$$

Therefore in the supply equation

$$Q_S = \alpha_1 + \alpha_2 P + U^S,$$

P is correlated with U^S. The OLS estimator is

$$\begin{aligned}
\hat{\alpha}_2^{OLS} =& \frac{\sum_{i=1}^n \left(P_i - \overline{P}\right)\left(Q_i - \overline{Q}\right)}{\sum_{i=1}^n \left(P_i - \overline{P}\right)^2} \\
=& \alpha_2 + \frac{\sum_{i=1}^n \left(P_i - \overline{P}\right)\left(U_i^S - \overline{U}^S\right)}{\sum_{i=1}^n \left(P_i - \overline{P}\right)^2} \\
\longrightarrow_p& \alpha_2 + \frac{\mathrm{Cov}\left(P_i, U_i^S\right)}{\mathrm{Var}\left(P_i\right)}
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Cov}\left(P_i, U_i^S\right) =& \mathrm{Cov}\left(\frac{1}{\alpha_2 - \beta_2}\left(\beta_1 - \alpha_1 + \beta_3 Y_i + \beta_4 POP_i + U_i^D - U_i^S\right), U_i^S\right) \\
=& -\frac{1}{\alpha_2 - \beta_2}\mathrm{Var}\left(U_i^S\right) \\
\neq& 0
\end{aligned}$$

assuming that Y and POP are exogenous and so $\mathrm{Cov}\left(U^S, Y\right) = \mathrm{Cov}\left(U^S, POP\right) = 0$. We are told that U^S and U^D are distributed independently, so that $\mathrm{Cov}\left(U^S, U^D\right) = 0$.

The instrument variable estimator is

$$\begin{aligned}
\hat{\alpha}_2^{IV} =& \frac{\sum_{i=1}^n \left(Y_i - \overline{Y}\right)\left(Q_i - \overline{Q}\right)}{\sum_{i=1}^n \left(Y_i - \overline{Y}\right)\left(P_i - \overline{P}\right)} \\
=& \alpha_2 + \frac{\sum_{i=1}^n \left(Y_i - \overline{Y}\right)\left(U^S - \overline{U}^S\right)}{\sum_{i=1}^n \left(Y_i - \overline{Y}\right)\left(P_i - \overline{P}\right)} \\
\longrightarrow_p& \alpha_2 + \frac{\mathrm{Cov}\left(Y_i, U_i^S\right)}{\mathrm{Cov}\left(P_i, Y_i\right)}.
\end{aligned}$$

The desired result follows from the assumptions $\mathrm{Cov}\left(Y_i, U_i^S\right) = 0$ and $\mathrm{Cov}\left(P_i, Y_i\right) \neq 0$.

Problem 4. Consider the simple regression model (with independently and identically distributed (i.i.d.) observations):

$$Y_i = \beta_0 + \beta_1 X_i^* + U_i.$$

Assume that $\mathbb{E}U_i = \mathbb{E}X_i^* U_i = 0$. However, instead of observing X_i^*, we only observed $X_i = X_i^* + e_i$. We think of X_i as some measurement of X_i^* that is subject to error. Assume

$$\mathbb{E}e_i = \mathbb{E}e_i U_i = \mathbb{E}X_i^* e_i = 0.$$

(i) Suppose we estimate the model using LS with the observed X_i in place of X_i^*. Let $\widehat{\beta}_1^{LS}$ denote the LS estimator. Show that

$$\widehat{\beta}_1^{LS} \to_p \beta_1 \frac{\mathrm{Var}\left(X_i^*\right)}{\mathrm{Var}\left(X_i^*\right) + \mathrm{Var}\left(e_i\right)}.$$

This means when there is measurement error, the LS estimate is closer to zero than β_1.

(ii) Suppose we have a second (subject-to-error) measurement of X_i^*, Z_i such that $\text{Cov}(Z_i, X_i^*) \neq 0$, $\mathbb{E}Z_i e_i = 0$ and $\mathbb{E}Z_i U_i = 0$. Show that

$$\widetilde{\beta}_1 = \frac{\sum_{i=1}^n \left(Z_i - \overline{Z}\right) Y_i}{\sum_{i=1}^n \left(Z_i - \overline{Z}\right) X_i}$$

is a consistent estimator for β_1. This means the classical measurement error problem can be resolved if two independent (subject-to-error) measurements of the same variable are available.

Solution. (i) The linear model:

$$Y_i = \beta_0 + \beta_1 X_i^* + U_i = Y_i = \beta_0 + \beta_1 (X_i - e_i) + U_i = \beta_0 + \beta_1 X_i + U_i - \beta_1 e_i.$$

The LS estimator:

$$\widehat{\beta}_1^{LS} = \beta_1 + \frac{n^{-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)(U_i - \beta_1 e_i)}{n^{-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}.$$

$\mathbb{E}(X_i U_i) = \mathbb{E}(X_i^* + e_i) U_i = 0$ and $\mathbb{E}(X_i e_i) = \mathbb{E}(X_i^* + e_i) e_i = \text{Var}(e_i)$. $\text{Var}(X_i) = \text{Var}(X_i^*) + \text{Var}(e_i) + 2 \cdot \text{Cov}(X_i^*, e_i) = \text{Var}(X_i^*) + \text{Var}(e_i)$. So, by WLLN and Continuous Mapping Theorem,

$$\begin{aligned}
\frac{n^{-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)(U_i - \beta_1 e_i)}{n^{-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} &= \frac{n^{-1}\sum_{i=1}^n \left(X_i - \overline{X}\right) U_i - \beta_1 n^{-1}\sum_{i=1}^n \left(X_i - \overline{X}\right) e_i}{n^{-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} \\
&\to_p \frac{\mathbb{E}(X_i U_i) - \mathbb{E}(X_i)\mathbb{E}(U_i) - \beta_1 \left(\mathbb{E}(X_i e_i) - \mathbb{E}(X_i)\mathbb{E}(e_i)\right)}{\text{Var}(X_i)} \\
&= -\beta_1 \frac{\text{Var}(e_i)}{\text{Var}(X_i^*) + \text{Var}(e_i)}.
\end{aligned}$$

(ii) Same as problem 2.

Problem 5. Consider the linear model (with independently and identically distributed (i.i.d.) observations):

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i$$

with $\mathbb{E}U_i = \mathbb{E}U_i X_{1,i} = \mathbb{E}U_i X_{2,i} = 0$. Suppose we know that $\beta_2 = \beta_1$ and conduct a constrained LS estimation of β_1:

$$\min_{b_0, b_1} \sum_{i=1}^n \left(Y_i - b_0 - b_1 X_{1,i} - b_1 X_{2,i}\right)^2.$$

(i) Find the expression for the constrained LS estimator $\left(\widehat{\beta}_0, \widehat{\beta}_1\right)$ that solve the above minimization problem.

(ii) Assume that the restriction $\beta_2 = \beta_1$ is true. Derive the large-sample (asymptotic) distribution of $\widehat{\beta}_1$.

Solution. Denote $\overline{X}_1 = n^{-1}\sum_{i=1}^n X_{1,i}$ and $\overline{X}_2 = n^{-1}\sum_{i=1}^n X_{2,i}$. The constrained LS:

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n \left(X_{1,i} + X_{2,i} - \overline{X}_1 - \overline{X}_2\right) Y_i}{\sum_{i=1}^n \left(X_{1,i} + X_{2,i} - \overline{X}_1 - \overline{X}_2\right)^2}.$$

And

$$\widehat{\beta}_0 = \overline{Y} - \widehat{\beta}_1 \left(\overline{X}_1 + \overline{X}_2\right).$$

For (ii),

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n \left(X_{1,i} + X_{2,i} - \overline{X}_1 - \overline{X}_2\right) Y_i}{\sum_{i=1}^n \left(X_{1,i} + X_{2,i} - \overline{X}_1 - \overline{X}_2\right)^2}$$

$$
\begin{aligned}
&= \frac{\sum_{i=1}^n \left(X_{1,i}+X_{2,i}-\overline{X}_1-\overline{X}_2\right)\left(\beta_0+\beta_1 X_{1,i}+\beta_1 X_{2,i}+U_i\right)}{\sum_{i=1}^n \left(X_{1,i}+X_{2,i}-\overline{X}_1-\overline{X}_2\right)^2} \\
&= \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^n \left(X_{1,i}+X_{2,i}-\overline{X}_1-\overline{X}_2\right)U_i}{\frac{1}{n}\sum_{i=1}^n \left(X_{1,i}+X_{2,i}-\overline{X}_1-\overline{X}_2\right)^2}.
\end{aligned}
$$

By WLLN and Continuous Mapping Theorem,

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^n \left(X_{1,i}+X_{2,i}-\overline{X}_1-\overline{X}_2\right)^2 &= \frac{1}{n}\sum_{i=1}^n \left(X_{1,i}+X_{2,i}\right)^2-\left(\overline{X}_1+\overline{X}_2\right)^2 \\
&\to_p \operatorname{Var}\left(X_{1,i}+X_{2,i}\right).
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^n \left(X_{1,i}+X_{2,i}-\overline{X}_1-\overline{X}_2\right)U_i &= \frac{1}{n}\sum_{i=1}^n \left(X_{1,i}+X_{2,i}-\mathbb{E}\left(X_{1,i}+X_{2,i}\right)\right)U_i \\
&\quad + \left(\overline{X}_1-\mathbb{E}\left(X_{1,i}\right)\right)\frac{1}{n}\sum_{i=1}^n U_i + \left(\overline{X}_2-\mathbb{E}\left(X_{2,i}\right)\right)\frac{1}{n}\sum_{i=1}^n U_i.
\end{aligned}
$$

Since $n^{-1/2}\sum_{i=1}^n U_i \to_d N\left(0,\mathbb{E}\left(U_i^2\right)\right)$, $\overline{X}_1-\mathbb{E}\left(X_{1,i}\right)\to_p 0$ and $\overline{X}_2-\mathbb{E}\left(X_{2,i}\right)\to_p 0$, by Slutsky's theorem,

$$
\begin{aligned}
\left(\overline{X}_1-\mathbb{E}\left(X_{1,i}\right)\right)\frac{1}{\sqrt{n}}\sum_{i=1}^n U_i &\to_p 0 \\
\left(\overline{X}_2-\mathbb{E}\left(X_{2,i}\right)\right)\frac{1}{\sqrt{n}}\sum_{i=1}^n U_i &\to_p 0.
\end{aligned}
$$

By CLT,

$$
\frac{1}{\sqrt{n}}\sum_{i=1}^n \left(X_{1,i}+X_{2,i}-\mathbb{E}\left(X_{1,i}+X_{2,i}\right)\right)U_i \to_d N\left(0,\mathbb{E}\left(U_i^2\left(X_{1,i}+X_{2,i}-\mathbb{E}\left(X_{1,i}+X_{2,i}\right)\right)^2\right)\right).
$$

By Slutsky's theorem,

$$
\begin{aligned}
\sqrt{n}\left(\widehat{\beta}_1-\beta_1\right) &= \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^n \left(X_{1,i}+X_{2,i}-\overline{X}_1-\overline{X}_2\right)U_i}{\frac{1}{n}\sum_{i=1}^n \left(X_{1,i}+X_{2,i}-\overline{X}_1-\overline{X}_2\right)^2} \\
&\to_d \operatorname{Var}\left(X_{1,i}+X_{2,i}\right)^{-1} N\left(0,\mathbb{E}\left(U_i^2\left(X_{1,i}+X_{2,i}-\mathbb{E}\left(X_{1,i}+X_{2,i}\right)\right)^2\right)\right).
\end{aligned}
$$

Problem 6. Suppose we observe the i.i.d. random sample $\{(Y_i, X_i)\}_{i=1}^n$ with X_i being a scalar. Take the linear model

$$
\begin{aligned}
Y_i &= X_i\beta + e_i \\
\mathbb{E}\left(e_i|X_i\right) &= 0.
\end{aligned}
$$

Consider the estimator

$$
\widehat{\beta} = \frac{\sum_{i=1}^n X_i^3 Y_i}{\sum_{i=1}^n X_i^4}.
$$

Find the asymptotic distribution of $\sqrt{n}\left(\widehat{\beta}-\beta\right)$.

Solution. The estimator:

$$
\widehat{\beta} = \frac{\sum_{i=1}^n X_i^3 Y_i}{\sum_{i=1}^n X_i^4} = \frac{\sum_{i=1}^n X_i^3\left(X_i\beta+e_i\right)}{\sum_{i=1}^n X_i^4} = \beta + \frac{n^{-1}\sum_{i=1}^n X_i^3 e_i}{n^{-1}\sum_{i=1}^n X_i^4}
$$

and

$$\sqrt{n}\left(\widehat{\beta}-\beta\right)=\frac{n^{-1/2}\sum_{i=1}^{n}X_i^3e_i}{n^{-1}\sum_{i=1}^{n}X_i^4}$$

By LIE, $\mathbb{E}\left(X_i^3e_i\right)=\mathbb{E}\left(X_i^3\mathbb{E}\left(e_i|X_i\right)\right)=0$. By WLLN, $n^{-1}\sum_{i=1}^{n}X_i^4\rightarrow_p\mathbb{E}\left(X_i^4\right)$. By CLT, $n^{-1/2}\sum_{i=1}^{n}X_i^3e_i\rightarrow_d N\left(0,\mathbb{E}\left(e_i^2X_i^6\right)\right)$. By Slutsky's theorem,

$$\sqrt{n}\left(\widehat{\beta}-\beta\right)\rightarrow_d\mathbb{E}\left(X_i^4\right)^{-1}N\left(0,\mathbb{E}\left(e_i^2X_i^6\right)\right)\sim N\left(0,\mathbb{E}\left(X_i^4\right)^{-2}\mathbb{E}\left(e_i^2X_i^6\right)\right).$$

Problem 7. Suppose we observe the i.i.d. random sample $\{(Y_i,X_i)\}_{i=1}^n$ with X_i being a scalar. Take the linear model

$$Y_i=X_i\beta+e_i$$
$$\mathbb{E}\left(e_i|X_i\right)=0.$$

The parameter of interest is $\theta=\beta^2$. Consider the LS estimate $\widehat{\beta}$ and $\widehat{\theta}=\widehat{\beta}^2$. $\boldsymbol{X}=(X_1,...,X_n)'$. Find $\mathbb{E}\left(\widehat{\theta}|\boldsymbol{X}\right)$ using our knowledge of $\mathbb{E}\left(\widehat{\beta}|\boldsymbol{X}\right)$ and $\boldsymbol{V}_{\widehat{\beta}}=\operatorname{Var}\left(\widehat{\beta}|\boldsymbol{X}\right)$.

Solution. $\mathbb{E}\left(\widehat{\theta}|\boldsymbol{X}\right)=\mathbb{E}\left(\widehat{\beta}^2|\boldsymbol{X}\right)=\operatorname{Var}\left(\widehat{\beta}|\boldsymbol{X}\right)+\mathbb{E}\left(\widehat{\beta}|\boldsymbol{X}\right)^2=\boldsymbol{V}_{\widehat{\beta}}+\beta^2$. So $\widehat{\theta}$ is biased.

Problem 8. Suppose we observe the i.i.d. random sample $\{(Y_i,X_i)\}_{i=1}^n$ with X_i being a scalar. Take the linear model

$$Y_i=X_i\beta+e_i$$
$$\mathbb{E}\left(e_i|X_i\right)=0$$
$$\Omega=\mathbb{E}\left(X_i^2e_i^2\right).$$

Let $\widehat{\beta}$ be the LS estimate of β with residuals $\widehat{e}_i=Y_i-X_i\widehat{\beta}$. Consider the estimates of Ω:

$$\widetilde{\Omega}=\frac{1}{n}\sum_{i=1}^{n}X_i^2e_i^2$$
$$\widehat{\Omega}=\frac{1}{n}\sum_{i=1}^{n}X_i^2\widehat{e}_i^2.$$

(a) Find the asymptotic distribution of $\sqrt{n}\left(\widetilde{\Omega}-\Omega\right)$. (b) Find the asymptotic distribution of $\sqrt{n}\left(\widehat{\Omega}-\Omega\right)$.

Solution. By CLT,

$$\sqrt{n}\left(\widetilde{\Omega}-\Omega\right)\rightarrow_d N\left(0,\operatorname{Var}\left(X_i^2e_i^2\right)\right).$$

$\operatorname{Var}\left(X_i^2e_i^2\right)=\mathbb{E}\left(X_i^4e_i^4\right)-\mathbb{E}\left(X_i^2e_i^2\right)^2$.

Use the expansion

$$\begin{aligned}\widehat{e}_i^2&=\left(Y_i-X_i\widehat{\beta}\right)^2\\&=\left(e_i-X_i\left(\widehat{\beta}-\beta\right)\right)^2\\&=e_i^2+X_i^2\left(\widehat{\beta}-\beta\right)^2-2e_iX_i\left(\widehat{\beta}-\beta\right).\end{aligned}$$

Then,

$$\widehat{\Omega}=\frac{1}{n}\sum_{i=1}^{n}X_i^2\widehat{e}_i^2$$

$$= \widetilde{\Omega} + \frac{1}{n}\sum_{i=1}^{n} X_i^4 \left(\widehat{\beta}-\beta\right)^2 - 2\frac{1}{n}\sum_{i=1}^{n} X_i^3 e_i \left(\widehat{\beta}-\beta\right)$$

and

$$\sqrt{n}\left(\widehat{\Omega}-\Omega\right) = \sqrt{n}\left(\widetilde{\Omega}-\Omega\right) + \frac{1}{n}\sum_{i=1}^{n} X_i^4 \sqrt{n}\left(\widehat{\beta}-\beta\right)^2 - 2\frac{1}{n}\sum_{i=1}^{n} X_i^3 e_i \sqrt{n}\left(\widehat{\beta}-\beta\right).$$

Note that we have

$$\sqrt{n}\left(\widehat{\beta}-\beta\right) \to_d N\left(0, V_\beta\right)$$
$$\frac{1}{n}\sum_{i=1}^{n} X_i^4 \to_p \mathbb{E}\left(X_i^4\right)$$
$$\frac{1}{n}\sum_{i=1}^{n} X_i^3 e_i \to_p \mathbb{E}\left(X_i^3 e_i\right) = 0$$
$$\widehat{\beta}-\beta \to_p 0.$$

By Slutsky's theorem,

$$\frac{1}{n}\sum_{i=1}^{n} X_i^4 \sqrt{n}\left(\widehat{\beta}-\beta\right)^2 \to_p 0$$
$$\frac{1}{n}\sum_{i=1}^{n} X_i^3 e_i \sqrt{n}\left(\widehat{\beta}-\beta\right) \to_p 0.$$

So

$$\frac{1}{n}\sum_{i=1}^{n} X_i^4 \sqrt{n}\left(\widehat{\beta}-\beta\right)^2 - 2\frac{1}{n}\sum_{i=1}^{n} X_i^3 e_i \sqrt{n}\left(\widehat{\beta}-\beta\right) \to_p 0$$

and by Slutsky's theorem,

$$\sqrt{n}\left(\widehat{\Omega}-\Omega\right) \to_d N\left(0, \operatorname{Var}\left(X_i^2 e_i^2\right)\right).$$

Problem 9. Suppose we observe the i.i.d. random sample $\{(Y_i, X_i)\}_{i=1}^n$ with X_i being a scalar. Define the conditional mean $m(X_i) = \mathbb{E}(Y_i|X_i)$. A researcher is interested in estimating the average derivative

$$\theta = \mathbb{E}\left(m'(X_i)\right).$$

Assume that the true conditional mean takes the form

$$m(x) = c_0 + c_1 x + c_2 x^2. \tag{1}$$

But this is not necessarily known by the researcher. Write the moments of X_i as $\mu_X = \mathbb{E}X_i$, $\sigma_X^2 = \operatorname{Var}(X_i)$ and $s_X = \mathbb{E}(X_i - \mu_X)^3$.

(i) Given (1), find an expression for θ in terms of c_0, c_1, c_2 and the moments of X_i.

(ii) Suppose the researcher estimates θ by linear LS. Regress Y_i on X_i:

$$Y_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_i + \widehat{e}_i$$

and then set $\widehat{\theta} = \widehat{\beta}_1$. Suppose that the true linear projection is

$$\mathcal{P}(Y_i|X_i) = \beta_0 + \beta_1 X_i.$$

$\left(\widehat{\beta}_0, \widehat{\beta}_1\right)$ are consistent estimators of (β_0, β_1). Find the difference $\beta_1 - \theta$ in terms of c_0, c_1, c_2 and the moments of X_i. Hint: $\mathbb{E}X_i^2 = \sigma_X^2 + \mu_X^2$ and $\mathbb{E}X_i^3 = s_X + 3\mu_X\sigma_X^2 + \mu_X^3$.

(iii) Now suppose that the researcher knows that the quadratic specification (1) is the correct conditional mean and estimates a quadratic regression by regressing Y_i on X_i and X_i^2 (with an intercept):

$$Y_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_i + \widehat{\beta}_2 X_i^2 + \widehat{e}_i.$$

For simplicity, assume that the researcher knows the mean μ_X. Provide an appropriate consistent estimator $\widehat{\theta}$ for θ and show its consistency. Is $\widehat{\theta}$ unbiased?

(iv) Find the asymptotic distribution of $\sqrt{n}\left(\widehat{\theta} - \theta\right)$. It is sufficient to write your answer in terms of the asymptotic covariance matrix of the LS estimator. Hint: use Delta Method.

(v) Now suppose that μ_X is unknown. Provide an appropriate estimator $\widehat{\theta}$ for θ. Show that $\widehat{\theta}$ is consistent. How would you find the asymptotic distribution of $\sqrt{n}\left(\widehat{\theta} - \theta\right)$? Hint: use Delta Method.

Solution. (i) Since $m'(x) = c_1 + 2c_2 x$, $\theta = \mathbb{E}(c_1 + 2c_2 X_i) = c_1 + 2c_2\mu_X$.

(ii) By the formula for the best linear predictor, we know that

$$\beta_1 = \frac{\operatorname{Cov}(X_i, Y_i)}{\operatorname{Var}(X_i)} = \frac{\operatorname{Cov}(m(X_i), X_i)}{\operatorname{Var}(X_i)} = \frac{\operatorname{Cov}\left(c_0 + c_1 X_i + c_2 X_i^2, X_i\right)}{\operatorname{Var}(X_i)} = c_1 + c_2 \frac{\operatorname{Cov}\left(X_i^2, X_i\right)}{\sigma_X^2}.$$

$$\operatorname{Cov}\left(X_i^2, X_i\right) = \mathbb{E}\left(X_i^3\right) - \mathbb{E}\left(X_i^2\right)\mathbb{E}(X_i) = s_X + 2\mu_X\sigma_X^2.$$

$$\beta_1 = c_1 + c_2 \frac{s_X + 2\mu_X\sigma_X^2}{\sigma_X^2}.$$

$$\beta_1 - \theta = c_2\left(\frac{s_X + 2\mu_X\sigma_X^2}{\sigma_X^2} - 2\mu_X\right) = c_2\frac{s_X}{\sigma_X^2}.$$

(iii) $\widehat{\theta} = \widehat{\beta}_1 + 2\widehat{\beta}_2\mu_X$. Since the LS estimator is unbiased and consistent, $\widehat{\theta}$ is unbiased and consistent. Note that $\widehat{\theta}$ is a linear function of $\left(\widehat{\beta}_1, \widehat{\beta}_2\right)$.

$$\mathbb{E}\left(\widehat{\theta}\right) = \mathbb{E}\left(\widehat{\beta}_1\right) + 2\mathbb{E}\left(\widehat{\beta}_2\right)\mu_X = c_1 + 2c_2\mu_X$$

and by Continuous Mapping Theorem,

$$\widehat{\theta} \to_p c_1 + 2c_2\mu_X.$$

(iv) The LS estimators are asymptotically normal: $\widehat{\boldsymbol{\beta}} = \left(\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2\right)'$

$$\sqrt{n}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \to_d N\left(\mathbf{0}, \boldsymbol{V}_{\boldsymbol{\beta}}\right),$$

where $\boldsymbol{V}_{\boldsymbol{\beta}} = \boldsymbol{Q}^{-1}\boldsymbol{\Omega}\boldsymbol{Q}^{-1}$, $\boldsymbol{Q} = \mathbb{E}\left(\boldsymbol{X}_i\boldsymbol{X}_i'\right)$, $\boldsymbol{\Omega} = \mathbb{E}\left(\boldsymbol{X}_i\boldsymbol{X}_i' e_i^2\right)$, $\boldsymbol{X}_i = \left(1, X_i, X_i^2\right)'$. Set $\boldsymbol{R} = (0, 1, 2\mu_X)'$. Note that $\theta = \boldsymbol{R}'\boldsymbol{\beta}$ and $\widehat{\theta} = \boldsymbol{R}'\widehat{\boldsymbol{\beta}}$. Thus, by Delta method,

$$\sqrt{n}\left(\widehat{\theta} - \theta\right) \to_d N\left(0, \boldsymbol{R}'\boldsymbol{V}_{\boldsymbol{\beta}}\boldsymbol{R}\right).$$

(v) $\widehat{\theta} = \widehat{\beta}_1 + 2\widehat{\beta}_2\overline{X}$, where $\overline{X} = n^{-1}\sum_{i=1}^n X_i$. Let $g(\beta_0, \beta_1, \beta_2, \mu) = \beta_1 + 2\beta_2\mu$. Now we can write $\widehat{\theta} = g\left(\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2, \overline{X}\right)$ and $\theta = g(c_0, c_1, c_2, \mu_X)$. By Continuous Mapping Theorem, since the function g is continuous, $\widehat{\theta}$ is consistent.

$$\sqrt{n}\begin{pmatrix} \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta} \\ \overline{X} - \mu_X \end{pmatrix} = \begin{pmatrix} \left(\frac{1}{n}\sum_{i=1}^n \boldsymbol{X}_i\boldsymbol{X}_i'\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \boldsymbol{X}_i e_i\right) \\ \frac{1}{\sqrt{n}}\sum_{i=1}^n (X_i - \mu_X) \end{pmatrix}$$

$$= \left(\begin{array}{cc} \left(\frac{1}{n}\sum_{i=1}^n \boldsymbol{X}_i\boldsymbol{X}_i'\right)^{-1} & \mathbf{0} \\ \mathbf{0}' & 1 \end{array} \right) \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\begin{array}{c} \boldsymbol{X}_i e_i \\ X_i - \mu_X \end{array} \right).$$

By CLT,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\begin{array}{c} \boldsymbol{X}_i e_i \\ X_i - \mu_X \end{array} \right) \to_d N\left(\mathbf{0}, \left(\begin{array}{cc} \mathbb{E}\left(\boldsymbol{X}_i\boldsymbol{X}_i' e_i^2\right) & \mathbb{E}\left(\boldsymbol{X}_i e_i \left(X_i - \mu_X\right)\right) \\ \mathbb{E}\left(\boldsymbol{X}_i' e_i \left(X_i - \mu_X\right)\right) & \mathbb{E}\left(\left(X_i - \mu_X\right)^2\right) \end{array} \right)\right).$$

Note that by LIE, $\mathbb{E}\left(\boldsymbol{X}_i e_i \left(X_i - \mu_X\right)\right) = \mathbb{E}\left(\mathbb{E}\left(e_i | X_i\right) \boldsymbol{X}_i \left(X_i - \mu_X\right)\right) = \mathbf{0}$. So,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\begin{array}{c} \boldsymbol{X}_i e_i \\ X_i - \mu_X \end{array} \right) \to_d N\left(\mathbf{0}, \left(\begin{array}{cc} \boldsymbol{\Omega} & \mathbf{0} \\ \mathbf{0}' & \sigma_X^2 \end{array} \right)\right).$$

Let $\boldsymbol{Q}_{\boldsymbol{XX}} = \mathbb{E}\left(\boldsymbol{X}_i\boldsymbol{X}_i'\right)$. By Slutsky's theorem,

$$\sqrt{n}\left(\begin{array}{c} \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta} \\ \overline{X} - \mu_X \end{array} \right) \to_d \left(\begin{array}{cc} \boldsymbol{Q}_{\boldsymbol{XX}}^{-1} & \mathbf{0} \\ \mathbf{0}' & 1 \end{array} \right) N\left(\mathbf{0}, \left(\begin{array}{cc} \boldsymbol{\Omega} & \mathbf{0} \\ \mathbf{0}' & \sigma_X^2 \end{array} \right)\right) \sim N\left(\mathbf{0}, \left(\begin{array}{cc} \boldsymbol{Q}_{\boldsymbol{XX}}^{-1} & \mathbf{0} \\ \mathbf{0}' & 1 \end{array} \right)\left(\begin{array}{cc} \boldsymbol{\Omega} & \mathbf{0} \\ \mathbf{0}' & \sigma_X^2 \end{array} \right)\left(\begin{array}{cc} \boldsymbol{Q}_{\boldsymbol{XX}}^{-1} & \mathbf{0} \\ \mathbf{0}' & 1 \end{array} \right)\right).$$

Note that

$$\left(\begin{array}{cc} \boldsymbol{Q}_{\boldsymbol{XX}}^{-1} & \mathbf{0} \\ \mathbf{0}' & 1 \end{array} \right)\left(\begin{array}{cc} \boldsymbol{\Omega} & \mathbf{0} \\ \mathbf{0}' & \sigma_X^2 \end{array} \right)\left(\begin{array}{cc} \boldsymbol{Q}_{\boldsymbol{XX}}^{-1} & \mathbf{0} \\ \mathbf{0}' & 1 \end{array} \right) = \left(\begin{array}{cc} \boldsymbol{Q}_{\boldsymbol{XX}}^{-1}\boldsymbol{\Omega}\boldsymbol{Q}_{\boldsymbol{XX}}^{-1} & \mathbf{0} \\ \mathbf{0}' & \sigma_X^2 \end{array} \right).$$

Let

$$\widetilde{\boldsymbol{R}} = \left(\begin{array}{c} \frac{\partial}{\partial \beta_0} g\left(c_0, c_1, c_2, \mu_X\right) \\ \frac{\partial}{\partial \beta_1} g\left(c_0, c_1, c_2, \mu_X\right) \\ \frac{\partial}{\partial \beta_2} g\left(c_0, c_1, c_2, \mu_X\right) \\ \frac{\partial}{\partial \mu} g\left(c_0, c_1, c_2, \mu_X\right) \end{array} \right) = \left(\begin{array}{c} 0 \\ 1 \\ 2\mu_X \\ 2c_2 \end{array} \right) = \left(\begin{array}{c} \boldsymbol{R} \\ 2c_2 \end{array} \right).$$

Then by Delta Method,

$$\sqrt{n}\left(\widehat{\theta} - \theta\right) \to_d N\left(0, V_\theta\right),$$

where

$$V_\theta = \widetilde{\boldsymbol{R}}' \left(\begin{array}{cc} \boldsymbol{Q}_{\boldsymbol{XX}}^{-1}\boldsymbol{\Omega}\boldsymbol{Q}_{\boldsymbol{XX}}^{-1} & \mathbf{0} \\ \mathbf{0}' & \sigma_X^2 \end{array} \right) \widetilde{\boldsymbol{R}} = \boldsymbol{R}'\boldsymbol{V}_{\boldsymbol{\beta}}\boldsymbol{R} + 4c_2^2\sigma_X^2.$$

Problem 10. Consider the following simple regression model:

$$Y_i = \alpha + \beta X_i + U_i.$$

Suppose the observations (Y_i, X_i), $i = 1, 2, ..., n$ are iid. Assume $\mathbb{E}|U_i| < \infty$, $\mathbb{E}|X_i| < \infty$ and $\mathbb{E}U_i = 0$. Let $\widetilde{\beta}_n$ be any consistent estimator of β (not necessarily the LS estimator). Define the following estimator for α:

$$\widetilde{\alpha}_n = \overline{Y}_n - \widetilde{\beta}_n \overline{X}_n,$$

where $\overline{Y}_n = n^{-1}\sum_{i=1}^n Y_i$ and $\overline{X}_n = n^{-1}\sum_{i=1}^n X_i$. Prove that $\widetilde{\alpha}_n$ is a consistent estimator of α.

Solution. Since $\overline{Y}_n = \alpha + \beta\overline{X}_n + n^{-1}\sum_{i=1}^n U_i$,

$$\widetilde{\alpha}_n - \alpha = \beta\overline{X}_n + n^{-1}\sum_{i=1}^n U_i - \widetilde{\beta}_n\overline{X}_n = \left(\beta - \widetilde{\beta}_n\right)\overline{X}_n + n^{-1}\sum_{i=1}^n U_i.$$

Then, by WLLN and Continuous Mapping Theorem,

$$\widetilde{\alpha}_n - \alpha \to_p 0 \cdot \mathrm{E}\left[X_i\right] + 0 = 0.$$

Problem 11. Consider the following regression model without a regressor:

$$Y_i = \alpha + U_i.$$

Suppose the observations Y_i, $i = 1, 2, ..., n$ are iid and $\mathbb{E}Y_i^2 < \infty$. Assume $\mathbb{E}U_i = 0$. What is the expression of the LS estimator $\widehat{\alpha}_n$? Show that $\sqrt{n}\left(\widehat{\alpha}_n - \alpha\right) \to_d N\left(0, V\right)$ and find V.

Solution. The LS estimator $\widehat{\alpha}_n$ solves

$$\min_a \sum_{i=1}^{n} \left(Y_i - a\right)^2.$$

So $\widehat{\alpha}_n = n^{-1} \sum_{i=1}^{n} Y_i$. $\mathrm{E}\left[Y_i\right] = \alpha + \mathrm{E}\left[U_i\right] = \alpha$. By CLT, $\sqrt{n}\left(\widehat{\alpha}_n - \alpha\right) \to_d N\left(0, V\right)$ with $V = \mathrm{Var}\left(Y_i\right)$.

Problem 12. Let $\{\theta_n : n \geq 1\}$ be a random sequence such that $\Pr\left(\theta_n = 0\right) = \left(n-1\right)/n$, and $\Pr\left(\theta_n = n^2\right) = 1/n$. Note that the only possible values for θ_n are zero and n^2.

(i) Show that $\lim_{n\to\infty} \mathbb{E}\theta_n = \infty$.

(ii) Does θ_n converge in probability to some limit? If yes, prove. If not, explain why.

Solution. (i) $\mathbb{E}\theta_n = 0 \cdot \left(n-1\right)/n + n^2 \cdot 1/n = n \to \infty$.
(ii) $\theta_n \to_p 0$, since for any $\epsilon > 0$, $\Pr\left(\left|\theta_n\right| > \epsilon\right) = \Pr\left(\theta_n = n^2\right) = 1/n \to 0$.