Solutions for Homework 7

Tong Zhou

Problems: 10.30, 11.15, 12.8, 12.10, 12.11.

PROBLEM 10.30

The parameter θ can be derived as

$$\theta = \frac{\beta_1}{\beta_2 + \beta_3 \mathrm{exp}/50}.$$

when exp = 10, it becomes

$$\theta = \frac{\beta_1}{\beta_2 + 0.2 \times \beta_3}.$$

To obtain its asymptotic standard error, the matrix

$$R = \frac{\partial r(\beta)}{\partial \beta} = \begin{pmatrix} \partial\theta/\partial\beta_1 \\ \partial\theta/\partial\beta_2 \\ \partial\theta/\partial\beta_3 \\ \partial\theta/\partial\beta_4 \end{pmatrix} = \begin{pmatrix} \frac{1}{\beta_2+\beta_3 \mathrm{exp}/50} \\ -\frac{\beta_1}{(\beta_2+\beta_3 \mathrm{exp}/50)^2} \\ -\frac{\beta_1 \mathrm{exp}}{(\beta_2+\beta_3 \mathrm{exp}/50)^2 50} \\ 0 \end{pmatrix}.$$

It then follows that

$$V_\theta = R' V_{\widehat{\beta}} R$$

and

$$s\left(\widehat{\theta}\right) = \sqrt{\widehat{R}' \widehat{V}_{\widehat{\beta}} \widehat{R}}.$$

Below table 1 summarizes standard errors and confidence intervals for $\widehat{\theta}$ using three bootstrapping methods. Code is provided on page 6.

Their discrepancies arise from the fact that in the asymptotic normality, we assume large sample theory works well so that normality condition is a good description of the behavior of $\widehat{\theta}$. However, if we suspect the sample size 99 is large enough, then asymptotic normality may not be a good approximation. Thus, a more general nonparametric bootstrapping technique might be more appropriate. Here we present three bootstrapping results: paired bootstrap, residual bootstrap and wild bootstrap. In practice, wild bootstrap is perhaps preferred over the other two, because it takes into account both conditional mean restriction and conditional heteroskedasticity. For more details on them, please consult Hansen's text.

Method	Standard Error	90% CI
Normality	0.870	(1.467, 4.330)
Paired bootstrap	1.304	(1.961, 4.908)
Residual bootstrap	1.528	(1.835, 5.161)
Wild bootstrap	1.143	(1.955, 4.754)

TABLE 1
Summary of results

PROBLEM 11.15

PROOF. (a) The appropriate estimators for $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are:

$$\widehat{\beta}_1 = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}_1$$
$$\widehat{\beta}_2 = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}_2$$

(b) The two regressions can be equivalently rewritten as

$$\begin{pmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{pmatrix} = \begin{pmatrix}\boldsymbol{X} & \boldsymbol{0}_{n\times k}\\ \boldsymbol{0}_{n\times k} & \boldsymbol{X}\end{pmatrix}\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} + \begin{pmatrix}\boldsymbol{e}_1\\ \boldsymbol{e}_2\end{pmatrix}.$$

Define

$$\mathcal{X} = \begin{pmatrix}\boldsymbol{X} & \boldsymbol{0}_{n\times k}\\ \boldsymbol{0}_{n\times k} & \boldsymbol{X}\end{pmatrix}, \quad \boldsymbol{y} = \begin{pmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{pmatrix}$$

and

$$\boldsymbol{e} = \begin{pmatrix}\boldsymbol{e}_1\\ \boldsymbol{e}_2\end{pmatrix}.$$

Then

$$\begin{pmatrix}\widehat{\beta}_1\\ \widehat{\beta}_2\end{pmatrix} = (\mathcal{X}'\mathcal{X})^{-1}\mathcal{X}'\boldsymbol{y}.$$

To derive its asymptotic distribution, first write

$$\begin{pmatrix}\widehat{\beta}_1\\ \widehat{\beta}_2\end{pmatrix} - \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} = (\mathcal{X}'\mathcal{X})^{-1}\mathcal{X}'\boldsymbol{e}.$$

Second, write $\mathcal{X}$ as an observation-by-observation form:

$$\mathcal{X} = \begin{pmatrix}\boldsymbol{x}_1'\\ \boldsymbol{x}_2'\\ \vdots\\ \boldsymbol{x}_{2n}'\end{pmatrix}$$

where each $\boldsymbol{x}_i$ is $2\boldsymbol{k}$-dimensional.
Then the joint asymptotic distribution of $\widehat{\beta}_1$ and $\widehat{\beta}_2$ is

$$\sqrt{n}\left(\begin{bmatrix}\widehat{\beta}_1\\ \widehat{\beta}_2\end{bmatrix} - \begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}\right) \xrightarrow{\mathsf{d}} \mathcal{N}\left(\boldsymbol{0}, Q_{\boldsymbol{xx}}^{-1}\Omega Q_{\boldsymbol{xx}}^{-1}\right) \tag{1}$$

where $Q_{\boldsymbol{xx}} = \mathbb{E}[\boldsymbol{x}_i\boldsymbol{x}_i']$ and $\Omega = \mathbb{E}[\boldsymbol{x}_i\boldsymbol{x}_i'e_i^2]$.

(c) To test $H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$, apply the matrix

$$R' = \begin{bmatrix} I_k & I_k \end{bmatrix}$$

to the asymptotic distribution of Eq. (1) and find a χ^2-distributed Wald statistics.

□

PROBLEM 12.8

PROOF. We propose three numerically equivalent methods.

Method 1: 2SLS

Take the demand function as an example.

- 1st Stage: Regress P on instrument variables $\begin{bmatrix} 1 & Y & W \end{bmatrix}$, and obtain the fitted value $\widehat{P}$.
- 2nd Stage: Regress Q on $\begin{bmatrix} 1 & \widehat{P} & Y \end{bmatrix}$

The supply function can be consistently estimated in the same way.

Method 2: IV Methods

Like method 1, all estimators can be consistently estimated separately. For each regression, since the numbers of instrument variables and regressors are equal, IV method can be directly employed by formula (12.26) of Hansen's notes:

$$\begin{pmatrix} \hat{a}_0 & \hat{a}_1 & \hat{a}_2 \end{pmatrix}' = \left(\boldsymbol{Z}'\boldsymbol{X}_1\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{Q}\right)$$
$$\begin{pmatrix} \hat{b}_0 & \hat{b}_1 & \hat{b}_2 \end{pmatrix}' = \left(\boldsymbol{Z}'\boldsymbol{X}_2\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{Q}\right),$$

where $\boldsymbol{X}_1$ includes $1, P$ and Y, $\boldsymbol{X}_2$ includes $1, P$ and W.

Method 3: Simultaneous Estimation

P and Q can be solved for:

$$\begin{cases} P &= \dfrac{b_0 - a_0}{a_1 - b_1} + \left(\dfrac{b_2}{a_1 - b_1}\right) W - \left(\dfrac{a_2}{a_1 - b_1}\right) Y + \dfrac{e_2 - e_1}{a_1 - b_1} \\ Q &= \dfrac{a_1 b_0 - a_0 b_1}{a_1 - b_1} + \left(\dfrac{a_1 b_2}{a_1 - b_1}\right) W - \left(\dfrac{a_2 b_1}{a_1 - b_1}\right) Y + \dfrac{a_1 e_2 - b_1 e_1}{a_1 - b_1}. \end{cases}$$

Observe that

$$\begin{cases} a_1 &= \dfrac{\mathsf{Cov}(Q, W)}{\mathsf{Cov}(P, W)} \\ \\ b_1 &= \dfrac{\mathsf{Cov}(Q, Y)}{\mathsf{Cov}(P, Y)} \end{cases} \tag{2}$$

So natural estimators for a_1 and b_1 are:

$$\begin{aligned}\widehat{a}_1 &= \frac{\text{sample covariance btw Q and W}}{\text{sample covariance btw P and W}} \\ &= \frac{\sum_i (Q_i - \overline{Q})(W_i - \overline{W})}{\sum_i (P_i - \overline{P})(W_i - \overline{W})}.\end{aligned}$$

and

$$\begin{aligned}\widehat{b}_1 &= \frac{\text{sample covariance btw Q and Y}}{\text{sample covariance btw P and Y}} \\ &= \frac{\sum_i (Q_i - \overline{Q})(Y_i - \overline{Y})}{\sum_i (P_i - \overline{P})(Y_i - \overline{Y})}.\end{aligned}$$

Then a_0 and a_2 can be estimated by a simple regression of $(Q - \widehat{a}_1 P)$ on Y, while b_0 and b_2 can be estimated by a simple regression of $(Q - \widehat{b}_1 P)$ on W.

It is easy to show all estimators are consistent.

□

PROBLEM 12.10

Proof. **(a)**

By

$$\mathbb{E}[\boldsymbol{x}_i \varepsilon_i] = \boldsymbol{\Gamma}' \underbrace{\mathbb{E}[\boldsymbol{z}_i e_i]}_{=\mathbf{0}} - \boldsymbol{\Gamma}' \underbrace{\mathbb{E}[\boldsymbol{z}_i \boldsymbol{u}_i']}_{=\mathbf{0}} \boldsymbol{\gamma} + \underbrace{\mathbb{E}[\boldsymbol{u}_i \varepsilon_i]}_{=\mathbf{0}} = \mathbf{0}$$

(b)

Define $\boldsymbol{v}_i = \begin{pmatrix} \boldsymbol{x}_i \\ \boldsymbol{u}_i \end{pmatrix}$, then

$$\mathbb{E}[\boldsymbol{v}_i \varepsilon_i] = \mathbf{0}$$

and

$$y_i = \boldsymbol{v}_i' \begin{pmatrix} \beta \\ \gamma \end{pmatrix} + \varepsilon_i.$$

Then, the OLS estimator

$$\begin{pmatrix} \widehat{\beta} \\ \widehat{\gamma} \end{pmatrix} = \begin{pmatrix} \beta \\ \gamma \end{pmatrix} + \left(\sum_i \boldsymbol{v}_i \boldsymbol{v}_i' \right)^{-1} \left(\sum_i \boldsymbol{v}_i \varepsilon_i \right).$$

Its asymptotic distribution can be routinely established:

$$\sqrt{n} \left(\begin{pmatrix} \widehat{\beta} \\ \widehat{\gamma} \end{pmatrix} - \begin{pmatrix} \beta \\ \gamma \end{pmatrix} \right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(\mathbf{0}, \boldsymbol{Q}_{vv}^{-1} \boldsymbol{\Omega} \boldsymbol{Q}_{vv}^{-1} \right)$$

where $\boldsymbol{Q}_{vv} = \mathbb{E}[\boldsymbol{v}_i \boldsymbol{v}_i']$ and $\boldsymbol{\Omega} = \mathbb{E}[\boldsymbol{v}_i \boldsymbol{v}_i' \varepsilon_i^2]$.

□

PROBLEM 12.11

PROOF. (a) Without further information, x^2 should be treated as an endogenous regressor.

(b) Define the matrix of instrument variables $\boldsymbol{Z}$, which includes $\left[1, z_i, z_i^2\right]$ with dimension $n \times 3$, and $\boldsymbol{X}$ includes all regressors 1, x, and x^2 with the same dimension. We know $\widehat{\beta} = (\boldsymbol{Z}'\boldsymbol{X})^{-1}\boldsymbol{Z}'\boldsymbol{y}$. Hence, the condition under which parameters are identified can be shown to be:

$$\det(\boldsymbol{Z}'\boldsymbol{X}) \neq 0.$$

It can be derived that this condition is equivalent to $\rho^2_{z,z^2} \neq 1$, where ρ_{z,z^2} is the correlation coefficient between z and z^2.

δ

(c) $x^2 = \gamma_0^2 + \gamma_1^2 z^2 + 2\gamma_0\gamma_1 z + u^2 + 2\gamma_0\gamma_1 z u + 2\gamma_0 u$. Therefore, if z and z^2 satisfy the relevance condition of an eligible IV, $\gamma_1 \neq 0$.

□

R CODE

```r

#### 10.30

rm(list=ls())
library(pacman)
p_load(tidyverse,haven,mpoly,corpcor,purrr,ggplot2,quantreg,tsibble,brolgar,viridisLite,
       viridis,plotly,magrittr,data.table,rio,jtools)

DT <- import(file = "https://www.ssc.wisc.edu/~bhansen/econometrics/cps09mar.dta",
    setclass = "data.table")

DT <- DT[hisp == 1 & female == 0 & race == 1 & marital == 7 & region == 2]

n = nrow(DT)

B <- 10000 ## number of replications

DT[, `:=`( lwage = log(earnings/(hours*week)), exper = age-education -6 )][,exper2
                                                                   := exper^2/100]
DT <- DT[,.(lwage,education,exper,exper2)]
fit <- DT %>% lm(lwage ~ education+exper+exper2,.)
summ(fit, digits = 4, confint = TRUE)
coef_dt <- DT[, .( coeff = coef( lm( lwage ~ education+exper+exper2)))] %>% as.matrix
var_dt <- DT[, .( coeff = vcov( lm( lwage ~ education+exper+exper2)))] %>% as.matrix
residual_dt <- DT[, .(res = resid(fit) ) ]
DT[, res_wild := resid(fit)]

RR <- matrix(NA,nrow = 4,ncol = 1)
RR[2,1] <- 1 / ( coef_dt[3,1] + coef_dt[4,1] * 0.2 )
RR[3,1] <- - coef_dt[2,1] / ( coef_dt[3,1] + coef_dt[4,1] * 0.2 )^2
RR[4,1] <- - 0.2 * coef_dt[2,1] / ( coef_dt[3,1] + coef_dt[4,1] * 0.2 )^2
RR[1,1] <- 0

###  asymptotic standard error
se_theta_ols <- sqrt(t(RR) %*% var_dt %*% RR)
print(se_theta_ols)

### value of theta
theta_calc =coef_dt[[2,1]]/( coef_dt[[3,1]] + 0.2 * coef_dt[[4,1]] )
print(theta_calc)

### 90% confidence interval using normality
CI_lower <- theta_calc - 1.645 * se_theta_ols
CI_upper <- theta_calc + 1.645 * se_theta_ols
print(CI_lower)
print(CI_upper)

### Three bootstrapping methods: paired bootstrap / residual bootstrap / wild bootstrap
theta_pair <- rep(NA,B)
theta_res <- rep(NA,B)
```

```r
theta_wild <- rep(NA,B)
V <- rep(NA,n)

ptm <- proc.time()
set.seed(1003)
for ( i in 1 : B){
  ## paired bootstrap
  DT_reg_pair <- DT[sample(.N, .N,replace=TRUE)]
  coeff_pair <- DT_reg_pair %>%
     lm(lwage ~ education+exper+exper2,.) %>%
     coef
  theta_pair[i] <- coeff_pair[2]/( coeff_pair[3] + 0.2 * coeff_pair[4])

  ## residual bootstrap
  res_boots  <- residual_dt[sample(.N, .N,replace=TRUE)]
  DT[, res := res_boots[[1]] ]
  DT[, lwage_res :=coef_dt[1,1]+coef_dt[2,1]*education+coef_dt[3,1]*exper+coef_dt[4,1]*exper2+res]
  coeff_res <- DT %>%
     lm( lwage_res ~ education+exper+exper2,.) %>%
     coef
  theta_res[i] <- coeff_res[2]/( coeff_res[3] + 0.2 * coeff_res[4])

  ## wild bootstrap
   V <- rnorm(n)
   DT[, `:=`( lwage_wild = coef_dt[1,1]+coef_dt[2,1]*education+coef_dt[3,1]*exper+coef_dt[4,1]*exper2+ V*res_wild )]
   coeff_wild <- DT %>%
     lm( lwage_wild ~ education+exper+exper2,.) %>%
     coef
   theta_wild[i] <- coeff_wild[2]/( coeff_wild[3] + 0.2 * coeff_wild[4])
}
proc.time()-ptm

## paired bootstrap: se and confidence interval
se_theta_boots_pair <- sqrt(mean( (theta_pair - mean(theta_pair))^2))
print(se_theta_boots_pair)
quantile(theta_pair,c(.05,.95))

## residual bootstrap: se and confidence interval
se_theta_boots_res <- sqrt(mean( (theta_res - mean(theta_res))^2))
print(se_theta_boots_res)
quantile(theta_res,c(.05,.95))

## wild bootstrap: se and confidence interval
se_theta_boots_wild <- sqrt(mean( (theta_wild - mean(theta_wild))^2))
print(se_theta_boots_wild)
quantile(theta_wild,c(.05,.95))
```