Solutions for Homework 3

Tong Zhou

Problems: 3.12, 3.13, 3.24, 3.26 4.7.

PROBLEM 3.12

Proof. No. Equation (3.52) is suffering from perfect multicollinearity. To remedy this, one has to drop one variable from D_1, D_2 or the intercept. Equation (3.53) drops the intercept term, while equation (3.54) drops the dummy variable D_2.

(a) Equations (3.53) and (3.54) are equivalent. Notice that $D_1 + D_2 = \mathbb{1}_n$. Plugging

$$D_2 = \mathbb{1}_n - D_1$$

into (3.53) yields

$$Y = \alpha_2 + D_1(\alpha_1 - \alpha_2) + e. \tag{1}$$

Compare Eq. (1) and equation (3.54), it is not hard to see

$$\begin{aligned} \mu &= \alpha_2 \\ \phi &= \alpha_1 - \alpha_2 \end{aligned}$$

(b) Observe that

$$\begin{aligned} \mathbb{1}_n' D_1 &= n_1 \\ \mathbb{1}_n' D_2 &= n_2 \end{aligned}$$

□

PROBLEM 3.13

Proof. (a) Let the annihilator matrix

$$\begin{aligned} M_2 &= I_n - D_2(D_2'D_2)^{-1}D_2' \\ &= I_n - \frac{1}{n_2} D_2 D_2' \end{aligned}$$

where I_n is the identity matrix. Apply M_2 to the regression function, then

$$M_2 y = M_2 D_1 \widehat{\gamma}_1 + M_2 \widehat{u}.$$

Thus

$$\begin{aligned} \widehat{\gamma}_1 &= (D_1' M_2 D_1)^{-1}(D_1' M_2 y) \\ &= (D_1' D_1)^{-1} D_1' y \\ &= \bar{Y}_1. \end{aligned}$$

Similarly, it can be shown that $\widehat{\gamma}_2 = \bar{Y}_2$.

Remark 1. *An alternative solution is via the fact that D_1 and D_2 are orthogonal, i.e. $D_1'D_2 = D_2'D_1 = 0$. Then there is no need to invoke the annihilator matrix, justified by the Frisch-Waugh-Lovell theorem.*

(b) The two equations can be seen as two residuals from the regressions of Y on (D_1, D_2) and X on (D_1, D_2) respectively.
(c) By the Frisch-Waugh-Lovell theorem, $\widetilde{\beta}$ and $\widehat{\beta}$ should be algebraically identically.

$\square$

R codes for problem 3.24 and 3.26 are provided below.

PROBLEM 3.24

(a) $R^2 = 0.389$ and $\mathsf{SSE} = 82.5$.
(b) $R^2 = 0.369$ and $\mathsf{SSE} = 82.5$. Slope estimates are equal: 0.144.
(c) SSE is the same, but R^2's differ.

PROBLEM 3.26

Characteristic	Beta	95% CI[1]	p-value
(Intercept)	1.208	1.108, 1.307	<0.001
education	0.087	0.082, 0.092	<0.001
exper	0.028	0.022, 0.033	<0.001
exper2	-0.036	-0.046, -0.026	<0.001
d_NE	0.063	-0.011, 0.138	0.10
d_S	-0.066	-0.127, -0.006	0.032
d_W	0.018	-0.042, 0.078	0.6
d_married	0.191	0.147, 0.235	<0.001
d_widow	0.091	0.009, 0.172	0.029
d_sep	0.020	-0.094, 0.133	0.7

[1] CI = Confidence Interval

Figure 1. *Regression result*

PROBLEM 4.7

PROOF. (a)

$$\begin{aligned}\mathbb{E}\left(\widetilde{\beta}_{\mathsf{gls}}\middle|\boldsymbol{X}\right) &= \left(\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\Sigma^{-1}\mathbb{E}(\boldsymbol{y}|\boldsymbol{X})\\ &= \left(\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}\beta\\ &= \beta.\end{aligned}$$

(b) The variance of the GLS estimator $\widetilde{\beta}_{\mathrm{gls}}$ is:

$$\begin{aligned}\mathsf{Var}\left(\widetilde{\beta}_{\mathsf{gls}}\middle|\boldsymbol{X}\right) &= \left(\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\Sigma^{-1}\mathsf{Var}\left(\boldsymbol{y}|\boldsymbol{X}\right)\Sigma^{-1}\boldsymbol{X}\left(\boldsymbol{X}\Sigma^{-1}\boldsymbol{X}\right)^{-1}\\ &= \left(\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\Sigma^{-1}\Omega\Sigma^{-1}\boldsymbol{X}\left(\boldsymbol{X}\Sigma^{-1}\boldsymbol{X}\right)^{-1}\\ &= c^2\left(\boldsymbol{X}\Sigma^{-1}\boldsymbol{X}\right)^{-1}\\ &= (\boldsymbol{X}\Omega^{-1}\boldsymbol{X})^{-1}\end{aligned}$$

(c) Since

$$\boldsymbol{M}_1\boldsymbol{y} = \underbrace{\boldsymbol{M}_1\boldsymbol{X}}_{=\mathbf{0}}\beta + \boldsymbol{M}_1\boldsymbol{e} = \boldsymbol{M}_1\boldsymbol{e},$$

and

$$\begin{aligned}\boldsymbol{M}_1\boldsymbol{y} &= (\boldsymbol{I} - \boldsymbol{X}\left(\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\Sigma^{-1})\boldsymbol{y}\\ &= \boldsymbol{y} - \boldsymbol{X}\underbrace{\left(\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\Sigma^{-1}\boldsymbol{y}}_{=\widetilde{\beta}}\\ &= \widetilde{\boldsymbol{e}},\end{aligned}$$

then $\boldsymbol{M}_1\boldsymbol{e} = \widetilde{\boldsymbol{e}}$.

(d) Plugging $\boldsymbol{M}_1$ of (c) into $\boldsymbol{M}_1'\Sigma^{-1}\boldsymbol{M}_1$.

(e)

$$
\begin{aligned}
\mathbb{E}\left(\tilde{c}^2|\boldsymbol{X}\right) &= \frac{1}{(n-k)}\mathbb{E}\left(\tilde{\boldsymbol{e}}'\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{e}}|\boldsymbol{X}\right)\\
&= \frac{1}{(n-k)}\mathbb{E}\left(\mathrm{tr}(\tilde{\boldsymbol{e}}'\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{e}})|\boldsymbol{X}\right) \quad (\tilde{\boldsymbol{e}}'\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{e}} \text{ is a scalar})\\
&= \frac{1}{(n-k)}\mathbb{E}\left(\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{e}}\tilde{\boldsymbol{e}}')|\boldsymbol{X}\right)\\
&= \frac{1}{(n-k)}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\mathbb{E}\left(\boldsymbol{M}_1\boldsymbol{e}\boldsymbol{e}'\boldsymbol{M}_1'|\boldsymbol{X}\right)\right)\\
&= \frac{1}{(n-k)}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{M}_1\mathbb{E}\left(\boldsymbol{e}\boldsymbol{e}'|\boldsymbol{X}\right)\boldsymbol{M}_1'\right)\\
&= \frac{1}{(n-k)}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{M}_1\Omega\boldsymbol{M}_1'\right)\\
&= \frac{1}{(n-k)}\mathrm{tr}\left(\Omega\boldsymbol{M}_1'\boldsymbol{\Sigma}^{-1}\boldsymbol{M}_1\right)\\
&= \frac{c^2}{(n-k)}\mathrm{tr}\left(\Sigma\,\underbrace{\boldsymbol{M}_1'\Sigma^{-1}\boldsymbol{M}_1}_{\text{use (d)}}\right)\\
&= \frac{c^2}{(n-k)}\mathrm{tr}\left(\Sigma\left(\Sigma^{-1}-\Sigma^{-1}\boldsymbol{X}\left(\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\Sigma^{-1}\right)\right)\\
&= \frac{c^2}{(n-k)}\left[\mathrm{tr}\left(\boldsymbol{I}_n\right)-\mathrm{tr}\left(\boldsymbol{X}\left(\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\Sigma^{-1}\right)\right]\\
&= \frac{c^2}{(n-k)}\left(n-\mathrm{tr}\left(\underbrace{\left(\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}\right)^{-1}\left(\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}\right)}_{=\boldsymbol{I}_k}\right)\right)\\
&= \frac{c^2}{(n-k)}(n-k)\\
&= c^2
\end{aligned}
$$

Remark 2. *We repeatedly use the property*

$$\mathrm{tr}(AB) = \mathrm{tr}(BA)$$

for any matrices A and B if both AB and BA are meaningful.

(f) In terms of unbiasedness, it is a reasonable estimator.

□

R CODES FOR 3.24 AND 3.26

```
rm(list=ls())

library(tidyverse)
library(knitr)
library(gtsummary)   ## Summarize regression results
library(moderndive)  ## Summarize regression results
library(haven)       ## Use function read_stata to load dataset.

### load data
df <- read_stata("~/Dropbox/2021_Spring/Econometrics/hw3/cps09mar.dta")

####################################################################
                  # Exercise 3.24 #
#####################################################################

####### (a) ##########

#### create new variables: log_wage, experience, experience2
df <- df %>% mutate(log_wage = log(earnings/(hours*week)),
                    exper = age - education - 6,
                    exper2 = exper^2/100)

#### subsample
df_subsample <- df %>% filter(race == 4, marital==7,female==0,exper < 45)

#### linear regression
model <- lm(log_wage~education+exper+exper2,data=df_subsample)

#### obtain R2 and SSE
get_regression_summaries(model)  %>%
  mutate(SSE = mse*nobs) %>%
        select(r_squared,SSE) %>%
        kable()

####### (b) ##########

### two regressions
model_1 <- lm(log_wage~exper+exper2,data=df_subsample)
model_2 <- lm(education~exper+exper2,data=df_subsample)

### obtain residuals
error_1 <- get_regression_points(model_1) %>% select(residual)
error_2 <- get_regression_points(model_2) %>% select(residual)

df_error <- bind_cols(error_1,error_2) %>% setNames(c("error1","error2"))

### regression residual_1 to residual_2
model <- lm(error1~error2,data=df_error)

### obtain slope estimates
```

```
model %>% tbl_regression(.,intercept=T,
                         estimate_fun = function(x) style_ratio(x,digit=3)
                            )

### obtain R2 and SSE
get_regression_summaries(model)%>%
  mutate(SSE = mse*nobs) %>%
  select(r_squared,SSE) %>%
  kable()

####################################################################
                  # Exercise 3.26 #
#####################################################################

#### create new dummy variables
df <- df %>%
     mutate(d_NE = ifelse(region==1,1,0),
            d_S = ifelse(region==3,1,0),
            d_W = ifelse(region==4,1,0),
            d_married = ifelse(marital==1 | marital==2,1,0),
            d_widow= ifelse(marital==4 | marital==5,1,0),
            d_sep = ifelse(marital==6,1,0)) %>%
            select(-age)

#### do regression and obtain table
df %>% filter(race==1,female==0,hisp==1) %>%
  lm(log_wage~education+exper+exper2+d_NE+d_S+
       d_W+d_married+d_widow+d_sep, data=.) %>%
  tbl_regression(.,intercept=T,
                 estimate_fun = function(x) style_ratio(x,digits=3),
                 label=list(exper~"exper",exper2~"exper2"))
```