Econometrics-Jeffrey M. Wooldridge / Chapter 2

Furkan Zengin

27 08 2021

The Simple Regression Model

$$y = \beta_0 + \beta_1 x + u \tag{2.1}$$

$$E(u) = 0 \tag{2.5}$$

$$E(u|x) = E(u) \tag{2.6}$$

$$E(y|x) = \beta_0 + \beta_1 x \tag{2.8}$$

Deriving the Ordinary Least Squares Estimates

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})((y_i - \bar{y}))}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Example 2.3

```
library(wooldridge)

View(ceosal1)

attach(ceosal1)

MODEL1=lm(salary~roe)

summary(MODEL1)

plot(roe,log(salary)) #take log for better appearance
```

Example 2.4

```
attach(wage1)
MODEL2=lm(wage~educ)
summary(MODEL2)

predict1=predict(MODEL2,interval = "confidence")
predict1 # shows that fitted values and lower and upper intervals
fitted.values(MODEL2) # shows the same

wage_hat=-0.90+0.54*c(2:8)
wage_hat
```

Example 2.5

```
attach(vote1)

MODEL3=lm(voteA~shareA)

summary(MODEL3)

plot(shareA,voteA)

abline(MODEL3,col="darkred",lwd=3)


# sample average of OLS residuals is zero

avols=sum(residuals(MODEL3))
avols # is nearly zero

#As a result independent variables*residuals is also zero
#The point (x_bar,y_bar) is always on OLS regression line

plot(shareA,voteA)
abline(MODEL3,col="darkred",lwd=3)

x_bar=mean(shareA)
y_bar=mean(voteA)
points(x_bar,y_bar,col="green",pch=15)
```

Residuals

$$\sum_{i=1}^{n} \hat{u}_i(\hat{y}_i - \bar{y}) = 0$$

```
# sample average of OLS residuals is zero

avols=sum(residuals(MODEL3))
avols # is nearly zero
```

Example 2.12

```
attach(meap93)

MODEL4 = lm(math10 ~ lnchprg)

summary(MODEL4)
```

Estimating the Error Variance

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{n} \hat{u}_i \tag{2.61}$$

$$se(\hat{\beta}_1) = \frac{\hat{\sigma}}{\sqrt{SST_x}} = \frac{\hat{\sigma}}{(\sum_{i=1}^{n}(x_i - \bar{x}))^{\frac{1}{2}}}$$

Computer Exercises -/- C3

```
attach(sleep75)

MODEL5 = lm(sleep ~ totwrk)

summary(MODEL5)

coefplot::coefplot(MODEL5)
```

Econometrics-Jeffrey M. Wooldridge / Chapter 3

Furkan Zengin

28 08 2021

Multiple Regression Analysis: Estimation

$$wage = \beta_0 + \beta_1 educ + \beta_2 exper + u \tag{3.1}$$

$$cons = \beta_0 + \beta_1 inc + \beta_2 inc^2 + u \tag{3.4}$$

$$\frac{\Delta cons}{\Delta inc} \approx \beta_1 + 2\beta_2 inc$$

The Model with k Independent Variables

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots \beta_k x_k + u \tag{3.6}$$

$$E[u|x_1, x_2, \dots, x_k] = 0 \tag{3.8}$$

Obtaining the OLS Estimates

Minimize:

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \cdots - \beta_k x_{ik})^2 \quad (3.12)$$

Example 3.1

```
library(wooldridge)

attach(gpa1)

MODEL1=lm(colGPA~hsGPA+ACT)

summary(MODEL1)

plot(hsGPA , colGPA, col = "steelblue",pch = 20)

abline(MODEL1)
```

Example 3.2

```
attach(wage1)

MODEL2=lm(lwage~educ+exper+tenure)

summary(MODEL2)

library(rgl)

par(mar=c(0,0,0,0))

plot3d(x=educ,y=exper,z=lwage,
 col = "red",type = "s",
 radius = 0.1)
```

A "Partialling Out" Interpretation of Multiple Regression

$$\frac{(\sum_{i=1}^{n} \hat{r}_{i1} y_i)}{(\sum_{i=1}^{n} \hat{r}_{i1}^2)} \tag{3.22}$$

where the r_{i1} are the OLS residuals from a simple regression of x1 on x2

```
MODEL3=lm(lwage~educ+exper)
summary(MODEL3)

MODEL4=lm(educ~exper)
resid1=residuals(MODEL4)

MODEL5=lm(lwage~resid1)
summary(MODEL5)
```

Goodness-of-Fit

$$R^2 = \frac{\sum_i (\hat{y}_i - \overline{y})^2}{\sum_i (y_i - \overline{y})^2} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \overline{y})^2}$$

Example 3.5

```
library(stargazer)

attach(crime1)

MODEL6 = lm(narr86~pcnv+ptime86+qemp86)

MODEL6_1=lm(narr86~pcnv+avgsen+ptime86+qemp86)

stargazer(MODEL6,MODEL6_1,type = "text")
```

Omitted Variable Bias: The Simple Case

	$\text{Cov}(x_1, x_2) > 0$	$\text{Cov}(x_1, x_2) < 0$
$\beta_2 > 0$	Positive Bias	Negative Bias
$\beta_2 < 0$	Negative Bias	Positive Bias

—

$$Var(\hat{\beta}_j) = \frac{\sigma^2}{SST_j}.VIF$$

```
cor1=cor(wage1)

#Omitting Variables causes bias if there is a correlation between independent variables

library(corrplot)

corrplot(cor1,type = "upper",order = "hclust",tl.col = "black", tl.srt = 45)

#High correlation bwt variables is considered as multicollinearity

car::vif(MODEL3)

car::vif(MODEL6_1)
```

Econometrics-Jeffrey M. Wooldridge / Chapter 4

Furkan Zengin

29 08 2021

Multiple Regression Analysis: Inference

Normality

$$\hat{\beta}_j \sim \mathcal{N}\left[\beta_j,\, Var\left(\beta_j\right)\right]$$

$$Where\ \ Var\left(\beta_j\right) = \frac{(\hat{\beta}_j - \beta_j)}{sd\left(\hat{\beta}_j\right)} \tag{4.1}$$

Null Hypothesis

$$H_o:\ \beta_j = 0 \tag{4.4}$$

```
devtools::install_github("cardiomoon/webr")

require(moonBook)

require(webr)

mean(wage)


t.test(wage1$wage,mu=5.896903)

plot(t.test(wage1$wage,mu=5.896903))

mean(wage1$educ)

plot(plot(t.test(wage1$educ,mu=12.56274)))
```

Example 4.1

```
library(wooldridge)

attach(wage1)

MODEL1=lm(lwage~educ+exper+tenure)

summary(MODEL1)

t_exper=0.004121/0.001723
t_exper #is greater than the critical value 1.645
```

Example 4.2

```
attach(meap93)

MODEL2=lm(math10~totcomp+staff+enroll)

summary(MODEL2)
```

Testing Other Hypotheses about β_j

Null Hypothesis

$$H_o: \ \beta_j = a_j \tag{4.12}$$

$$t = \frac{(\hat{\beta}_j - a_j)}{se\,(\hat{\beta}_j)}$$

$$t = \frac{(\text{estimate} - \text{hypothesized value})}{\text{standard error}} \tag{4.13}$$

Example 4.4

```
attach(campus)

MODEL3 = lm(lcrime ~ lenroll)

summary(MODEL3)

t = (MODEL3$coefficients[2] - 1)/0.1098
t
```

Example 4.5

```
attach(hprice2)

MODEL3=lm(lprice~log(nox)+log(dist)+rooms+stratio)

summary(MODEL3)
```

Computing p-Values for t Tests

$$\begin{aligned} p\text{-value} &= \Pr_{H_0}\left[\left|\frac{\hat{\beta}_j - \beta_j}{se\left(\hat{\beta}_j\right)}\right| > \left|\frac{\hat{\beta}_j^{act} - \beta_j}{se(\hat{\beta}_j}\right|\right] \\ &= \Pr_{H_0}\left(|t| > |t^{act}|\right) \\ &\approx 2 \cdot \Phi\left(-|t^{act}|\right) \end{aligned}$$

By using : https://www.econometrics-with-r.org/

```
#For the Example 4.5

library(scales)


t <- seq(-12, 12, 0.01)

plot(x = t,
     y = dnorm(t, 0, 1),
     type = "l",
     col = "steelblue",
     lwd = 2,
     yaxs = "i",
     axes = F,
     ylab = "",
     main = expression("Calculating the p-value of a Two-sided Test when" ~ t^act ~ "=11.567"),
     cex.lab = 0.7,
     cex.main = 1)

tact <- 11.567

axis(1, at = c(0, -1.96, 1.96, -tact, tact), cex.axis = 0.7)

# Shade the critical regions using polygon():

# critical region in left tail
polygon(x = c(-6, seq(-6, -1.96, 0.01), -1.96),
        y = c(0, dnorm(seq(-6, -1.96, 0.01)), 0),
        col = 'orange')
```

```r
# critical region in right tail

polygon(x = c(1.96, seq(1.96, 6, 0.01), 6),
        y = c(0, dnorm(seq(1.96, 6, 0.01)), 0),
        col = 'orange')

# Add arrows and texts indicating critical regions and the p-value
arrows(-3.5, 0.2, -2.5, 0.02, length = 0.1)
arrows(3.5, 0.2, 2.5, 0.02, length = 0.1)

arrows(-5, 0.16, -11.567, 0, length = 0.1)
arrows(5, 0.16, 11.567, 0, length = 0.1)

text(-3.5, 0.22,
     labels = expression("0.025"~"="~over(alpha, 2)),
     cex = 0.7)
text(3.5, 0.22,
     labels = expression("0.025"~"="~over(alpha, 2)),
     cex = 0.7)

text(-5, 0.18,
     labels = expression(paste("-|",t[act],"|")),
     cex = 0.7)
text(5, 0.18,
     labels = expression(paste("|",t[act],"|")),
     cex = 0.7)

# Add ticks indicating critical values at the 0.05-level, t^act and -t^act
rug(c(-1.96, 1.96), ticksize  = 0.145, lwd = 2, col = "darkred")
```

```
rug(c(-tact, tact), ticksize  = -0.0451, lwd = 2, col = "darkgreen")
```

Confidence Intervals

$$\hat{\beta}_j \pm c \cdot se\left(\hat{\beta}_j\right)$$

Example 4.8

```
attach(rdchem)

MODEL4=lm(log(rd)~log(sales)+profmarg)
summary(MODEL4)

confint(MODEL4,level = 0.95)

library(coefplot)

MODEL1=lm(lwage~educ+exper+tenure)

coefplot(MODEL4)
```

Testing Multiple Linear Restrictions: The F Test

$$F = \frac{\left(SSR_r - SSR_{ur}\right)/q}{SSR_{ur}/\left(n-k-1\right)} \tag{4.37}$$

```
attach(mlb1)

MODEL5=lm(log(salary)~years+gamesyr+bavg+hrunsyr+rbisyr)

summary(MODEL5)

#We will test whether bavg,hrunsyr,rbisyr have an effect on salary or not

library(car)

H0=c("bavg","hrunsyr","rbisyr")

linearHypothesis(MODEL5,H0)

#If we exclude bavg & hrunsyr from the model

MODEL6=lm(log(salary)~years+gamesyr+rbisyr)
summary(MODEL6)

#rbisyr is now statistically significant,that is why we cannot reject Null Hypothesis
```

Testing General Linear Restrictions

$$H_0 = \beta_1 = \beta_2 = \cdots = \beta_k \tag{4.44}$$

```r
attach(hprice1)

MODEL7=lm(log(price)~log(assess)+log(lotsize)+log(sqrft)+bdrms)

summary(MODEL7)

H01=c("assess=1","lotsize=0","sqrft=0","bdrms=0")

linearHypothesis(MODEL7,c("log(assess)=1","log(lotsize)=0","log(sqrft)=0","bdrms=0"))

library(stargazer)

MODEL8=lm(log(price)~log(assess)+log(lotsize))

MODEL9=lm(log(price)~log(assess)+log(lotsize)+log(sqrft))

MODEL10=lm(log(price)~log(assess)+log(lotsize)+log(sqrft)+bdrms)

stargazer(list(MODEL8,MODEL9,MODEL10),type = "text")
```

Econometrics-Jeffrey M. Wooldridge / Chapter 6

Furkan Zengin

30 08 2021

Multiple Regression Analysis: Further Issues

Effects of Data Scaling on OLS Statistics

$$\hat{bwght} = \hat{\beta}_0 + (20\hat{\beta}_1)(cigs/20) + \hat{\beta}_2 faminc$$

```
library(wooldridge)

attach(bwght)

MODEL1=lm(bwght~cigs+faminc)

summary(MODEL1)


MODEL2=lm(bwghtlbs~cigs+faminc)

summary(MODEL2)


packts=cigs/20

MODEL3=lm(bwght~packts+faminc)
summary(MODEL3)

library(stargazer)

stargazer(list(MODEL1,MODEL2,MODEL3),type = "text")
```

More on Functional Form

```
attach(hprice2)

MODEL4=lm(log(price)~log(nox)+rooms)

summary(MODEL4)

f1 = function(b){(100*(exp(b)-1))}

f1(0.306)
```

Models with Quadratic

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + u \tag{6.10}$$

```
attach(wage1)

MODEL5=lm(wage~exper)

MODEL6=lm(wage~exper+expersq)

summary(MODEL6)

plot(exper,wage,
 col="steelblue",pch=20)

abline(MODEL5,
 col="red", lwd=2)

ORDER1=order(exper)

lines(x=exper[ORDER1],
 y=fitted(MODEL6)[ORDER1],
 col="green",lwd=2)

x_star=abs(0.2981001/(2*-0.0061299))
x_star

y=3.72+0.30*24.31525-0.0061299*(24.31525)^2 #is the maximum value

points(24.31525,7.390386,lwd=3,col="red",pch=19)
```

Example 6.2

```
attach(hprice2)

roomssq=(rooms)^2

MODEL7=lm(log(price)~log(nox)+log(dist)+rooms+I(rooms^2)+stratio)

summary(MODEL7)

plot(rooms,log(price))

library(effects)

plot(effect("rooms",MODEL7))



f = data.frame(rooms=seq(4,8),nox=5.5498,dist=3.7958,stratio=18.4593)


p1 = predict(MODEL7, f, interval = "confidence")


cbind(f,p1)


matplot(f$rooms, p1, type="l", lty=c(1,2,2))
```

Models with Interaction Terms

$$y = \beta_0 + \beta_1 x + \beta_2 x2 + \beta_3 x_1 x_2 + u$$

Example 6.3

```
attach(attend)

data("attend")

priGPAsq=(priGPA)^2

ACTsq=(ACT)^2

MODEL8=lm(stndfnl~atndrte+priGPA+ACT+priGPAsq+ACTsq+priGPA*atndrte)

summary(MODEL8)
```

Prediction and Residual Analysis

```
options(scipen = 999)

attach(gpa2)

hsizesq=(hsize)^2

MODEL9=lm(colgpa~sat+hsperc+hsize+I(hsize^2),data = gpa2)

summary(MODEL9)

pred1=data.frame(sat=c(800,1000,1200,1400),hsperc=c(10,25,35,50),hsize=c(1,2,
3,5))

pred1

predict(MODEL9,pred1,interval = "confidence")
```

Econometrics-Jeffrey M. Wooldridge / Chapter 7

Furkan Zengin

31 08 2021

Multiple Regression Analysis with Qualitative Information

$$wage = \beta_0 + \delta_0 female + \beta_1 educ + u \tag{7.1}$$

$$\delta_0 = E(wage|female = 1, educ) - E(wage|female = 0, educ)$$

Example 7.1

```
library(wooldridge)

attach(wage1)

MODEL1=lm(lwage~female+educ+exper+tenure)

summary(MODEL1)

plot(lwage,educ, )

plot(educ,lwage, col = female + 1, pch = female + 1)
legend("bottomright", c("Male", "Female"), col = c(1, 2), pch = c(1, 2))


int1 = coef(MODEL1)[1]
int2 = coef(MODEL1)[1] + coef(MODEL1)[2]

slope1 = coef(MODEL1)[3]
slope2 = coef(MODEL1)[3]

abline(int1, slope1, col = 1, lty = 1, lwd = 2)
abline(int2, slope2, col = 2, lty = 2, lwd = 2)
```

Example 7.2

```
attach(gpa1)

MODEL2=lm(colGPA~PC+hsGPA+ACT)

summary(MODEL2)
```

Example 7.4

```
attach(hprice1)

MODEL3=lm(log(price)~log(lotsize)+log(sqrft)+bdrms+colonial)

summary(MODEL3)
```

Example 7.5

```
library(ggplot2)

MODEL4=lm(log(wage)~female+educ+exper+expersq+tenure+tenursq)

summary(MODEL4)

#We can see how getting married or being female affects wages below
#blue points refer to married and for columns 0 is male 1 is female

library(ggplot2)
ggplot(wage1,aes(educ,wage,col=factor(married)))+
  geom_point()+facet_wrap(~female)+geom_smooth(method = "lm")
```

Using Dummy Variables for Multiple Categories

Example 7.6

```
marrmale=ifelse(wage1$female==0 & wage1$married==1,1,0)

marrfem=ifelse(wage1$female==1 & wage1$married==1,1,0)

singfem=ifelse(wage1$female==1 & wage1$married==0,1,0)

#We create 3 dummy variables and set Married female as base group

MODEL5=lm(lwage~marrmale+marrfem+singfem+educ+exper+expersq+tenure+tenursq)

summary(MODEL5)
```

Example 7.8

```
attach(lawsch85)

)
r61_100=ifelse(lawsch85$rank > 60 & rank < 100,1,0)

MODEL6=lm(log(salary)~top10+r11_25+r26_40+r41_60+r61_100+
            LSAT+GPA+log(libvol+log(cost)))

summary(MODEL6)
```

Interactions Involving Dummy Variables

Example 7.10

```
attach(wage1)

MODEL7=lm(log(wage)~female+educ+female*educ+exper+expersq+tenure+tenursq)

summary(MODEL7)

# Section 7-4c Testing for Differences in Regression Functions across groups
attach(gpa3)

MODEL8=lm(cumgpa ~female*(sat+hsperc+tothrs),subset = (spring==1))
summary(MODEL8)

library(car)

H0=c("female","female:sat","female:hsperc","female:tothrs")
linearHypothesis(MODEL8,H0)
```

Econometrics-Jeffrey M. Wooldridge / Chapter 8

Furkan Zengin

01 09 2021

Heteroskedasticity

Homoscedasticity

$$Var(u_i|x_i) = \sigma_i^2$$

$$E(u_i^2) = \sigma^2 \tag{11.1.1}$$

Simulations of Heteroskedasticity Belong to Edward Rubin:
https://github.com/edrubin/EC421S20/tree/master/notes

```
options(htmltools.dir.version = FALSE)
library(pacman)
p_load(
  broom, latex2exp, ggplot2, ggthemes, viridis, extrafont,
  dplyr,
  magrittr, knitr, parallel
)


# Define pink color
red_pink <- "#e64173"
grey_light <- "grey70"
grey_mid <- "grey50"
grey_dark <- "grey20"
# Dark slate grey: #314f4f
# Notes directory
dir_slides <- "~/Dropbox/UO/Teaching/EC421W19/LectureNotes/02Review/"
# Knitr options
opts_chunk$set(
  comment = "#>",
  fig.align = "center",
  fig.height = 7,
  fig.width = 10.5,
  warning = F,
  message = F
)
# A blank theme for ggplot
theme_empty <- theme_bw() + theme(
  line = element_blank(),
  rect = element_blank(),
  strip.text = element_blank(),
  axis.text = element_blank(),
  plot.title = element_blank(),
  axis.title = element_blank(),
```

```
  plot.margin = structure(c(0, 0, -0.5, -1), unit = "lines", valid.unit = 3L, class = "unit"),
  legend.position = "none"
)
theme_simple <- theme_bw() + theme(
  line = element_blank(),
  panel.grid = element_blank(),
  rect = element_blank(),
  strip.text = element_blank(),
  axis.text.x = element_text(size = 18, family = "STIXGeneral"),
  axis.text.y = element_blank(),
  axis.ticks = element_blank(),
  plot.title = element_blank(),
  axis.title = element_blank(),
  # plot.margin = structure(c(0, 0, -1, -1), unit = "lines", valid.unit = 3L, class = "unit"),
  legend.position = "none"
)
theme_axes_math <- theme_void() + theme(
  text = element_text(family = "MathJax_Math"),
  axis.title = element_text(size = 22),
  axis.title.x = element_text(hjust = .95, margin = margin(0.15, 0, 0, 0, unit = "lines")),
  axis.title.y = element_text(vjust = .95, margin = margin(0, 0.15, 0, 0, unit = "lines")),
  axis.line = element_line(
    color = "grey70",
    size = 0.25,
    arrow = arrow(angle = 30, length = unit(0.15, "inches")
  )),
  plot.margin = structure(c(1, 0, 1, 0), unit = "lines", valid.unit = 3L, class = "unit"),
  legend.position = "none"
)
theme_axes_serif <- theme_void() + theme(
  text = element_text(family = "MathJax_Main"),
  axis.title = element_text(size = 22),
  axis.title.x = element_text(hjust = .95, margin = margin(0.15, 0, 0, 0, unit = "lines")),
  axis.title.y = element_text(vjust = .95, margin = margin(0, 0.15, 0, 0, unit = "lines")),
  axis.line = element_line(
```

```
    color = "grey70",
    size = 0.25,
    arrow = arrow(angle = 30, length = unit(0.15, "inches")
  )),
  plot.margin = structure(c(1, 0, 1, 0), unit = "lines", valid.unit = 3L, class = "unit"),
  legend.position = "none"
)
theme_axes <- theme_void() + theme(
  text = element_text(family = "Fira Sans Book"),
  axis.title = element_text(size = 18),
  axis.title.x = element_text(hjust = .95, margin = margin(0.15, 0, 0, 0, unit = "lines")),
  axis.title.y = element_text(vjust = .95, margin = margin(0, 0.15, 0, 0, unit = "lines")),
  axis.line = element_line(
    color = grey_light,
    size = 0.25,
    arrow = arrow(angle = 30, length = unit(0.15, "inches")
  )),
  plot.margin = structure(c(1, 0, 1, 0), unit = "lines", valid.unit = 3L, class = "unit"),
  legend.position = "none"
)
```

Examples of Heteroskedasticity

```
gq_df <- tibble(
  x = runif(1e3, -3, 3),
  e = rnorm(1e3, 0, sd = 4 + 1.5 * x),
  y = 1 + 3 * x + e)
```

```r
gq_x <- quantile(gq_df$x, probs = c(3/8, 5/8))
# Regressions
sse1 <- lm(y ~ x, data = gq_df %>% filter(x < gq_x[1])) %>%
  residuals() %>% raise_to_power(2) %>% sum()
sse2 <- lm(y ~ x, data = gq_df %>% filter(x > gq_x[2])) %>%
  residuals() %>% raise_to_power(2) %>% sum()
ggplot(data = gq_df, aes(x = x, y = e)) +
geom_point(color = "darkslategrey", size = 2.75, alpha = 0.5) +
labs(x = "x", y = "u") +
theme_axes_math
```

```r
ggplot(data = gq_df, aes(
  x = x, y = e,
  color = cut(x, c(-Inf, gq_x, Inf)),
  alpha = cut(x, c(-Inf, gq_x, Inf)),
  shape = cut(x, c(-Inf, gq_x, Inf))
)) +
geom_vline(
  xintercept = gq_x,
  color = grey_mid,
  size = 0.25
) +
geom_point(size = 2.75) +
labs(x = "x", y = "u") +
scale_color_manual(values = c("darkslategrey", grey_mid, red_pink)) +
scale_shape_manual(values = c(19, 1, 19)) +
scale_alpha_manual(values = c(0.5, 0.8, 0.6)) +
theme_axes_math
```

```r
set.seed(12345)
# Data
gq2_df <- tibble(
  x = runif(1e3, -3, 3),
  e = rnorm(1e3, 0, sd = 2 + x^2),
  y = 1 + 3 * x + e
)
# Quantiles
gq_x <- quantile(gq2_df$x, probs = c(3/8, 5/8))
# Regressions
sse1b <- lm(y ~ x, data = gq2_df %>% filter(x < gq_x[1])) %>%
  residuals() %>% raise_to_power(2) %>% sum()
sse2b <- lm(y ~ x, data = gq2_df %>% filter(x > gq_x[2])) %>%
  residuals() %>% raise_to_power(2) %>% sum()
ggplot(data = gq2_df, aes(
  x = x, y = e,
  color = cut(x, c(-Inf, gq_x, Inf)),
  alpha = cut(x, c(-Inf, gq_x, Inf)),
  shape = cut(x, c(-Inf, gq_x, Inf))
)) +
geom_vline(
  xintercept = gq_x,
  color = grey_mid,
  size = 0.25
) +
geom_point(size = 2.75) +
labs(x = "x", y = "u") +
scale_color_manual(values = c("darkslategrey", grey_mid, red_pink)) +
scale_shape_manual(values = c(19, 1, 19)) +
scale_alpha_manual(values = c(0.5, 0.8, 0.6)) +
theme_axes_math
```

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n}(x_i - \bar{x})u_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$Var(\hat{\beta}_1) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2\sigma_i^2}{SST_x^2} \tag{8.2}$$

Example 8.1

```
library(wooldridge)


library(lmtest)


marrmale=ifelse(wage1$female==0 & wage1$married==1,1,0)


marrfem=ifelse(wage1$female==1 & wage1$married==1,1,0)

singfem=ifelse(wage1$female==1 & wage1$married==0,1,0)

#We create 3 dummy variables and set Married female as base group

attach(wage1)
MODEL1=lm(lwage~marrmale+marrfem+singfem+educ+exper+
            expersq+tenure+tenursq)

summary(MODEL1)

library(sandwich)

robust=vcovHC(MODEL1,type = "HC1")

coeftest(MODEL1,robust)
```

```
attach(gpa3)

MODEL2=lm(cumgpa~sat+hsperc+tothrs+female+black+white,
          subset =(term==2))

summary(MODEL2)

robust2=vcovHC(MODEL2,type = "HC1")

coeftest(MODEL2,robust2)

H0=c("black","white")

library(car)

linearHypothesis(MODEL2,H0)
```

Testing for Heteroskedasticity

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k \tag{8.10}$$

$$H_0 : Var(u|x_1, x_2, \ldots, x_k) = \sigma^2 \tag{8.11}$$

```
attach(hprice1)

MODEL4=lm(price~lotsize+sqrft+bdrms)

bptest(MODEL4)

#For White Test

MODEL5=lm(log(price)~log(lotsize)+log(sqrft)+bdrms)

bptest(MODEL5,~fitted(MODEL5)+I(fitted(MODEL5)^2))
```

Weighted Least Squares Estimation

$$Var(u|\mathbf{x}) = \sigma^2 h(\mathbf{x}) \tag{8.21}$$

$$y = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + u_i \tag{8.24}$$

$$\frac{y_i}{\sqrt{h_i}} = \frac{\beta_0}{\sqrt{h_i}} + \beta_1 \frac{x_{1i}}{\sqrt{h_i}} + \beta_2 \frac{x_{2i}}{\sqrt{h_i}} + \frac{u_i}{\sqrt{h_i}} \tag{8.25}$$

```
library(wooldridge)

attach(k401ksubs)

#Using OLs

MODEL6=lm(nettfa~inc+I((age-25)^2)+male+e401k,subset=(fsize==1))

summary(MODEL6)

#Using WLS

MODEL7=lm(nettfa~inc+I((age-25)^2)+male+e401k,subset=(fsize==1),weights =
1/inc)

summary(MODEL7)

library(stargazer)

stargazer(MODEL6,MODEL7,type="text",column.labels = c("OLS","WLS"))

plot(inc,nettfa)

abline(MODEL6,lwd=3,lty=1,col="red")

abline(MODEL7,lwd=3,lty=3,col="green")

legend("topleft",c("OLS","WLS"),lty =c(1,3),bty = "n")
```

```
attach(smoke)

#Standard OLS

MODEL9=lm(cigs~lincome+lcigpric+educ+age+agesq+restaurn)

summary(MODEL9)

#Now we need residuals of previous regression

weight1=lm(log(residuals(MODEL9)^2)~lincome+lcigpric+educ+age+agesq+restaurn,
data = smoke)

MODEL10=lm(cigs~lincome+lcigpric+educ+age+agesq+restaurn,weights =
1/exp(fitted(weight1)),data = smoke)


summary(MODEL10)
stargazer(MODEL9,MODEL10,type="text",column.labels = c("OLS","WLS"))


#If we consider weight that we have chosen is wrong
#Then we can look at the Robust WLS.

library(lmtest)

library(sandwich)

robust3=vcovHC(MODEL7,type = "HC1")
MODEL7_Robust=coeftest(MODEL7,robust3)

stargazer(MODEL7,MODEL7_Robust,type = "text"
          ,column.labels = c("Non-Robust WLS","Robust WLS"))
```

$$Var(y|\mathbf{x}) = p(\mathbf{x})[1 - p(\mathbf{x})] \tag{8.45}$$

```
attach(mroz)

inlf2=ifelse(inlf==1,"Yes","No")

newmroz=data.frame(inlf2
 ,mroz)

View(newmroz)
MODEL11=lm(inlf~nwifeinc+educ+exper+expersq+age+kidslt6+kidsge6)

summary(MODEL11)


#We see the relationship between family income and the probability of participation

plot(nwifeinc,inlf,ylim = c(-0.4,1.5),main = "Family Income & Labour
Participation",xlab = "Family Income",ylab =
"Participation",pch=20,col="steelblue")

abline(h=1,lty=2,col="red")

abline(h=0,lty=2,col="red")

abline(MODEL11,lwd=2.2,col="green")



coeftest(MODEL11, vcov. = vcovHC, type = "HC1")
```

```
#We can determine weight for propability model

attach(gpa1)


parcoll=ifelse(mothcoll==1 | fathcoll==1,1,0)

newgpa1=data.frame(parcoll,gpa1)

View(newgpa1)
attach(newgpa1)

MODEL13=lm(PC~hsGPA+ACT+parcoll)

summary(MODEL13)

weight2=fitted(MODEL13)*(1-fitted(MODEL13))

MODEL14=lm(PC~hsGPA+ACT+parcoll,weights = 1/weight2)

summary(MODEL14)

stargazer(MODEL13,MODEL14,type = "text",column.labels = c("OLS","WLS"))
```

Econometrics-Jeffrey M. Wooldridge / Chapter 9

Furkan Zengin

02 09 2021

More on Specification and Data Issues

Functional Form Misspecification

$$log(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 exper^2 + \beta_4 female -$$

Example 9.1

```
library(wooldridge)


attach(crime1)

MODEL1=lm(narr86~pcnv+avgsen+tottime+ptime86+qemp86+inc86+black+hispan)

summary(MODEL1)

MODEL2=lm(narr86 ~ pcnv+ avgsen+ tottime+ ptime86+qemp86+
            inc86+ black+ hispan+ pcnvsq+ pt86sq+inc86sq)

stargazer::stargazer(list(MODEL1,MODEL2),type = "text")
```

RESET as a General Test for Functional Form Misspecification

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \delta_1 \hat{y}^2 + \delta_2 \hat{y}^3 + error \quad (9.3)$$

Example 9.2

```
attach(hprice1)

MODEL3=lm(price~lotsize+sqrft+bdrms)

MODEL4=lm(log(price)~llotsize+lsqrft+bdrms)

library(fRegression)

lmTest(price~lotsize+sqrft+bdrms,method = "reset")

lmTest(MODEL4,method = "reset")
```

Using Proxy Variables for Unobserved Explanatory Variables

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3^* + u \tag{9.10}$$

$$x_3^* = \delta_0 + \delta_3 x_3 + v_3 \tag{9.11}$$

$$E(x_3^* | x_1, x_2, x_3) = \delta_0 + \delta_3 x_3 \tag{9.13}$$

Example 9.3

```
attach(wage2)

View(wage2)

MODEL5=lm(lwage~educ+exper+tenure+married+south+urban+black)

MODEL6=lm(lwage~educ+exper+tenure+married+south+urban+black+IQ)

MODEL7=lm(lwage~educ+exper+tenure+married+south+urban+black+IQ+IQ*educ)

stargazer::stargazer(list(MODEL5,MODEL6,MODEL7),type = "text")
```

Example 9.4

```
library(magrittr)

attach(crime2)

library(dplyr)

tt87=crime2 %>% filter(year==87)

tt82=crime2 %>% filter(year==82)

MODEL8=lm(log(crmrte)~unem+log(lawexpc),data = tt87)

summary(MODEL8)

MODEL9=lm(log(tt87$crmrte)~tt87$unem+log(tt87$lawexpc)+log(tt82$crmrte))

summary(MODEL9)

stargazer::stargazer(list(MODEL8,MODEL9),type = "text")
```

Outliers and Influential Observations

```r
attach(rdchem)

MODEL10=lm(rdintens~sales+profmarg)

summary(MODEL10)

library(plotrix)

plot(sales,rdintens,col="steelblue",pch=20)

draw.circle(39709,3.5,radius = 1001, border="red",lty=3,lwd=3)


newrdchem=rdchem[-10,]


attach(newrdchem)

MODEL11=lm(rdintens~sales+profmarg)

summary(MODEL11)
```

Econometrics-Jeffrey M. Wooldridge / Chapter 10

Furkan Zengin

03 09 2021

Basic Regression Analysis with Time Series Data

Static Models

$$y_t = \beta_0 + \beta_1 z_t + u_t, \quad t = 1, 2, 3, \ldots, n \tag{10.1}$$

Finite Distributed Lag Models

$$y_t = \alpha_0 + \delta_1 z_t + \delta_2 z_{t-1} + \delta_3 z_{t-2} + u_t, \tag{10.5}$$

Example 10.1

```
library(wooldridge)

library(lmtest)

plot(unem,type = "l")

attach(phillips)

MODEL1=lm(inf~unem)

summary(MODEL1)


TSP <- ts(phillips, start=1948, frequency = 1)
inf <- ts(phillips$inf, start=1948, frequency = 1)
unem <- ts(phillips$unem, start=1948, frequency = 1)


ts.plot(unem,inf, col = c("blue", "red"), lty=2:3, xlab="Unemployment and Inflation")
legend("topright", bty="n", lty=1:2, col=c("blue","red"),
       legend=c(" Unemployment ", " Inflation "))
```

Example 10.2

```
attach(intdef)

MODEL2=lm(i3~inf+def)

summary(MODEL2)
View(intdef)

ti3 = ts(intdef$i3, start=1948, frequency = 1)

tinf = ts(intdef$inf, start=1948, frequency = 1)


ts.plot(tinf,ti3, col = c("blue", "red"), lty=2:3, xlab="Interest and Inflation")
legend("topright", bty="n", lty=1:2, col=c("blue","red"),
       legend=c(" Interest ", " Inflation "))
```


Example 10.3

```
attach(prminwge)

MODEL3=lm(log(prepop)~log(mincov)+log(usgnp))

summary(MODEL3)
```

Example 10.4

```
attach(fertil3)

MODEL4=lm(gfr~pe+ww2+pill)

summary(MODEL4)


MODEL5=lm(gfr~pe+ww2+pill+pe_1+pe_2)
summary(MODEL5)

# In order to find its signifance we create new variables.

a=pe_1-pe

b=pe_2-pe

MODEL6=lm(gfr~pe+a+b+ww2+pill)

summary(MODEL6)
```

Example 10.5

```
attach(barium)

MODEL7=lm(log(chnimp)~log(chempi)+log(gas)+
            log(rtwex)+befile6+affile6+afdec6)

summary(MODEL7)
```

Example 10.7

```
attach(hseinv)



MODEL8=lm(log(invpc)~log(price))

summary(MODEL8)

MODEL9=lm(log(invpc)~log(price)+t)

summary(MODEL9
```

Example 10.8

```
attach(fertil3)

MODEL10=lm(gfr~pe+ww2+pill+t)

summary(MODEL10)
```

Example 10.11

```
library(dynlm)

attach(barium)

seas=ts(barium,start = c(1978,2),frequency = 12)

MODEL11=dynlm(log(chnimp)~log(chempi)+log(gas)+
                log(rtwex)+befile6+affile6+afdec6+season(seas),data=seas)

summary(MODEL11)

library(car)

linearHypothesis(MODEL11,matchCoefs(MODEL11,"season"))
```

Econometrics-Jeffrey M. Wooldridge / Chapter 11

Furkan Zengin

04 09 2021

Further Issues in Using OLS with Time Series Data

Stationary and Weakly Dependent Time Series

Moving Average Process of Order One

$$x_t = e_t + \alpha_1 e_{t-1}, \quad t = 1, 2, \ldots, \tag{11.1}$$

Autoregressive Average Process of Order One

$$y_t = \rho_1 y_{t-1} + e_t, \quad t = 1, 2, \ldots, \tag{11.2}$$

$$Corr(y_t, y_{t+h}) = \rho_1^h \tag{11.4}$$

Asymptotic Properties of OLS

Zero Conditional Mean

$$\mathbf{x}_t = (x_{t1}, x_{t2}, \ldots x_{tk}):$$

$$E(u_t|\mathbf{x}_t) = 0$$

Homoskedasticity

$$Var(u_t|\mathbf{x}_t) = \sigma^2$$

No Serial Correlation

$$Var(u_t u_s|\mathbf{x}_t, \mathbf{x}_s) = 0$$

Example 11.4

```
library(wooldridge)

library(dynlm)

library(lmtest)

attach(nyse)


MODEL1=dynlm(return~return_1)

summary(MODEL1)

MODEL2=dynlm(ts(return)~L(ts(return))+L(ts(return,2)),data = nyse)

summary(MODEL2)
```

Example 11.5

```
library(car)

linearHypothesis(MODEL2, matchCoefs(MODEL2,"L"))

attach(phillips)

MODEL3=dynlm(cinf~unem)

summary(MODEL3)
```

Using Highly Persistent Time Series in Regression Analysis

Simulation code belongs to Hüseyin Taştan: https://github.com/htastan/Econometrics-II

```
library(forecast)
library(ggplot2)

set.seed(1234)
e1 <- ts(rnorm(500,0,1))
RW1 <- ts(cumsum(e1))
e2 <- ts(rnorm(500,0,1))
RW2 <- ts(cumsum(e2))
e3 <- ts(rnorm(500,0,1))
RW3 <- ts(cumsum(e3))
autoplot(RW1, main = "3 Independent Random Walk Process") +
  autolayer(RW2) + autolayer(RW3) + theme(legend.position = "none")
```

```
set.seed(1234)
# define the lists for the ARIMA(p,d,q) models
# order = c(1, 0, 0) means ARIMA(1,0,0) = AR(1)
# ar is the AR coefficient and sd is the standard deviation
list1 <- list(order = c(1, 0, 0), ar = 0.5, sd = 1)
list2 <- list(order = c(1, 0, 0), ar = 0.8, sd = 1)
list3 <- list(order = c(1, 0, 0), ar = 0.9, sd = 1)
list4 <- list(order = c(1, 0, 0), ar = 0.95, sd = 1)
#
AR1_1 <- arima.sim(n = 500, model = list1)
AR1_2 <- arima.sim(n = 500, model = list2)
AR1_3 <- arima.sim(n = 500, model = list3)
AR1_4 <- arima.sim(n = 500, model = list4)
#autoplot(AR1_1)

plot1 <- autoplot(AR1_1) + xlab("") + ggtitle("AR(1) = 0.5")
plot2 <- autoplot(AR1_2) + xlab("") + ggtitle("AR(1) = 0.8")
plot3 <- autoplot(AR1_3) + xlab("") + ggtitle("AR(1) = 0.9")
plot4 <- autoplot(AR1_4) + xlab("") + ggtitle("AR(1) = 0.95")
library(grid)
library(gridExtra)
grid.arrange(grobs=list(plot1, plot2, plot3, plot4),
             ncol=2, top="Simulated AR(1) Processes")
```

Transformations on Highly Persistent Time Series

$$\Delta y_t = y_t - y_{t-1} = e_t \tag{11.24}$$

Deciding Whether a Time Series Is I(1)

Example 11.6

```
attach(fertil3)

MODEL4=dynlm(cgfr~cpe)

summary(MODEL4)


MODEL5=dynlm(cgfr~cpe+cpe_1+cpe_2)
summary(MODEL5)

linearHypothesis(MODEL5,c("cpe=0","cpe_1=0"))
```

Example 11.7

```
attach(earns)

MODEL6=dynlm(lhrwage~loutphr+t)

summary(MODEL6)

MODEL7=lm(lhrwage~t)
MODEL8=lm(loutphr~t)

acf(residuals(MODEL7),1,pl=FALSE)

MODEL9=dynlm(ghrwage~goutphr)
summary(MODEL9)



MODEL10=dynlm(cgfr~cpe+cpe_1+cpe_2+cgfr_1)

summary(MODEL10) # So the last variable is significant
```

Econometrics-Jeffrey M. Wooldridge / Chapter 12

Furkan Zengin

05 09 2021

Serial Correlation and Heteroskedasticity in Time Series Regressions

---

Efficiency and Inference

$$u_t = \rho u_{t-1} + e_t \qquad (12.1)$$
$$|\rho| < 1$$

Serial Correlation–Robust Inference after OLS

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \cdots + \beta_k x_{kt} \qquad (12.10)$$

$$x_{t1} = \delta_0 + \delta_2 x_{t2} + \delta_3 x_{t3} + \cdots + \delta_k x_{kt} + r_t$$

$$\hat{v}_t = \sum_{t=1}^{n} \hat{a}_t^2 + 2\sum_{h=1}^{g}[1-(h/(g+1))](\sum_{t=h+1}^{n} \hat{a}_t\hat{a}_{t-h}) \tag{12.12}$$

Where $\hat{a}_t = \hat{r}_t\hat{u}_t$

$$se(\hat{\beta}_1) = [se(\hat{\beta}_1)/\hat{\sigma}]^2\sqrt{\hat{v}} \tag{12.13}$$

Example 12.1

```
library(wooldridge)

library(dynlm)

library(lmtest)

library(sandwich)


attach(prminwge)


MODEL1=dynlm(lprepop~lmincov+lusgnp+lprgnp+t)

summary(MODEL1)

coeftest(MODEL1,vcov=vcovHAC)
```

Example 12.2

```
attach(phillips)

data1=ts(phillips,start = 1948)

MODEL2=dynlm(inf~unem,data = data1,end = 2003)

summary(MODEL2)

res1=resid(MODEL2)

MODEL3=dynlm(res1 ~ L(res1))

summary(MODEL3)

MODEL4=dynlm(cinf~cunem,data = data1,end = 2003)

summary(MODEL4)

res2=resid(MODEL4)

MODEL5=dynlm(res2~L(res2))

summary(MODEL5)

dwtest(MODEL2)
dwtest(MODEL4)
```

Example 12.3

```
data2=ts(prminwge,start = 1950)

MODEL1rep=dynlm(lprepop~lmincov+lusgnp+lprgnp+t,data = data2,end = 1987)

res3=resid(MODEL1rep)

MODEL6=dynlm(res3~L(res3)+lmincov+lusgnp+lprgnp+t)

summary(MODEL6)
```

Example 12.4

``{r}
attach(barium)
data3=ts(barium,start = c(1978,2),frequency = 12)
MODEL7=dynlm(lchnimp~lchempi+lgas+lrtwex+befile6+ affile6+afdec6,data = data3)
res4=resid(MODEL7)
MODEL8=dynlm(res4~L(res4)+L(res4,2)+L(res4,3)+ lchempi+lgas+lrtwex+befile6+affile6+afdec6,data = data3)

library(car)
linearHypothesis(MODEL8,matchCoefs(MODEL8,“res”))
coeftest(MODEL7,vcov. = NeweyWest(MODEL7, lag = 4,adjust = TRUE,verbose = TRUE))

```

## Example 12.5

```
MODEL9=dynlm(lchnimp~lchempi+lgas+lrtwex+
              befile6+affile6+afdec6,data = data3)

library(prais)

MODEL10=prais_winsten(chnimp~lchempi+lgas+
             lrtwex+befile6+affile6+afdec6,data = data3)

library(stargazer)

prais_winsten(MODEL9,data = data3)

stargazer(list(MODEL9,MODEL10),type = "text",keep.stat = "n")

summary(MODEL10,object = summary.prais)
```
```

Econometrics-Jeffrey M. Wooldridge / Chapter 13

Furkan Zengin

06 09 2021

Pooling Cross Sections across Time: Simple Panel Data Methods

---

Example 13.1

```
library(wooldridge)

library(dynlm)

library(lmtest)

library(sandwich)

attach(fertil1)

library(stargazer)

library(plm)

MODEL1=lm(kids~educ+age+agesq+black+east+northcen+
            west+farm+othrural+town+
            smcity+y74+y76+y78+y80+y82+y84)
summary(MODEL1)

library(car)

h0=c("y82","y84")

linearHypothesis(MODEL1,h0)
```

Example 13.2

```
attach(cps78_85)

MODEL2=lm(lwage~y85+educ+y85educ+exper+I((exper^2)/100)+
            union+female+y85fem,data = cps78_85)

summary(MODEL2)

CPS78 = subset(cps78_85, year == "78")


CPS85 = subset(cps78_85, year == "85")
MODEL2_1=lm(lwage~y85+educ+y85educ+exper+I((exper^2)/100)+
              union+female+y85fem,data = CPS78)

MODEL2_2=lm(lwage~y85+educ+y85educ+exper+I((exper^2)/100)+
              union+female+y85fem,data = CPS85)

stargazer(list(MODEL2_1,MODEL2_2),type = "text")

library(ggplot2)

ggplot(cps78_85,aes(educ,lwage))+geom_point()+
  geom_smooth(method = "lm")+facet_wrap(~year)
```

Example 13.3

```
attach(kielmc)

kiel81=subset(kielmc,year=="1981")

MODEL3=lm(rprice~nearinc,data = kiel81)

summary(MODEL3)

MODEL4=lm(rprice~nearinc,subset = c(kielmc$year=="1978"))

summary(MODEL4)

MODEL5=lm(rprice~y81+y81nrinc+nearinc)

summary(MODEL5
```

Two-Period Panel Data Analysis

$$y_{it} = \beta_0 + \delta_0 d2_t + \beta_1 x_{it} + a_i + u_{it} \quad (13.20)$$

Example 13.5

```
attach(slp75_81)

pdata.frame(slp75_81)


MODEL6=lm(cslpnap~ctotwrk+ceduc+cmarr+cyngkid+cgdhlth)

summary(MODEL6)

h0_1=c("ceduc","cmarr","cyngkid","cgdhlth")

linearHypothesis(MODEL6,h0_1)
```

Example 13.6

```
attach(crime3)

MODEL7=lm(clcrime~cclrprc1+cclrprc2)

summary(MODEL7)
```

Example 13.8

```
attach(ezunem)



library(dynlm)

ezu=pdata.frame(ezunem,index = 22)

MODEL8=plm(diff(log(uclms))~diff(d81)+diff(d82)+
             diff(d83)+diff(d84)+diff(d85)
+diff(d86)+diff(d87)+diff(d88)+diff(ez),data = ezu)

summary(MODEL8)
```

Example 13.9

```
attach(crime4)

cri4=pdata.frame(crime4,index =90)



MODEL9=plm(diff(lcrmrte)~d83+d84+d85+d86+d87+
             diff(lprbarr)+
             diff(lprbconv)+diff(lprbpris)+
             diff(lavgsen)+diff(lpolpc),data = cri4)
summary(MODEL9)
```

Econometrics-Jeffrey M. Wooldridge / Chapter 18

Furkan Zengin

07 09 2021

Advanced Time Series Topics

Testing for Unit Roots

$$y_t = \alpha + \rho y_{t-1} + e_t \tag{18.17}$$

$$E(e_t | y_{t-1}, y_{t-2}, \ldots, y_0) = 0 \tag{18.18}$$

$$H_0 : \rho = 1 \tag{18.19}$$

$$H_1 : \rho < 1 \tag{18.20}$$

$$\Delta y_t = \alpha + \theta y_{t-1} + e_t \tag{18.21}$$

Example 18.2

```
library(wooldridge)

library(tseries)

attach(intqrt)

library(dynlm)


adf.test(intqrt$r3)

intq=ts(intqrt)

MODEL1=dynlm(diff(r3)~L(r3),data = intq)

summary(MODEL1)
```

$$\Delta y_t = \alpha + \theta y_{t-1} + \gamma_1 \Delta y_{t-1} + e_t \tag{18.23}$$

Example 18.3

```
attach(phillips)
adf.test(inf,k=1)
```

Example 18.4

```
library(wooldridge)

data(inven)

attach(inven)

inven$lgdp=log(inven$gdp)

invents=ts(inven)

adf.test(lgdp,k=1)

MODEL2=dynlm(diff(lgdp)~trend(invents)+L(diff(lgdp))+L(lgdp),data = invents)

summary(MODEL2)
```

Cointegration

```
attach(intqrt)

intq=ts(intqrt)


plot(merge(as.zoo(r3), as.zoo(r6)),
 plot.type = "single",
 lty = c(2, 1),
 lwd = 2,xlab = "Time",
 ylab = "Interest")

lines(as.zoo(spr63),
 col = "steelblue",
 lwd = 2)
polygon(c(time(r3), rev(time(r3))),c(r6, rev(r3)),col = "red")
legend("topleft",
 legend = c("r3", "r6", "spr63"),
 col = c("black", "red", "steelblue"),
 lwd = c(2, 2, 2),
 lty = c(2, 1, 1))


library(urca)
ur.df(window(r6) - window(r3),
 lags = 0,
 selectlags = "AIC",
```

```
type = "drift")
```

Example 18.5

```
attach(fertil3)

MODEL3=lm(gfr~pe)

res1=residuals(MODEL3)

adf.test(res1,k=1)

MODEL4=lm(gfr~pe+t)


res2=residuals(MODEL4)

adf.test(res2,k=1)
```

```
#Error Correction Models **Gujarati, Page 764**

MODEL6=lm(r6~r3)

res3=residuals(MODEL6)

MODEL7=dynlm(cr6~cr3+L(res3),data = intq)

summary(MODEL7)
```

Forecasting

```
library(forecast)

attach(phillips)

philts=ts(phillips,start = 1948)

MODEL8=dynlm(unem~unem_1,data = philts)

MODEL9=dynlm(unem~unem_1+inf_1,data = philts)

library(stargazer)

stargazer(list(MODEL8,MODEL9),type = "text")

pre1=predict(MODEL8,newdata = window(philts,start=1997),interval =
"prediction")

pre2=predict(MODEL9,newdata = window(philts,start=1997),interval =
"prediction")

jpre=data.frame(pre1,pre2)

forecast1=forecast(unem,h=6)

plot(forecast1)

accuracy(MODEL8)
```

```
accuracy(MODEL9)
```