Lab 12

AML

08/11/2020

```
setwd('/Users/andrea/Desktop/UEA/Classes/Econometrics/Data')
library(data.table)
df.OLSdat <- read.table("OLSdata.RData")
DT <- data.table(df.OLSdat)
summary(df.OLSdat)
```

```
##        y1                 const          x1                 x2          
##  Min.   :-11.7353   Min.   :1   Min.   :-1.0813748   Min.   :-3.626672  
##  1st Qu.: -0.2118   1st Qu.:1   1st Qu.:-0.2214267   1st Qu.:-0.686488  
##  Median :  2.2171   Median :1   Median : 0.0001322   Median : 0.000400  
##  Mean   :  2.1698   Mean   :1   Mean   :-0.0038033   Mean   :-0.006917  
##  3rd Qu.:  4.5833   3rd Qu.:1   3rd Qu.: 0.2136286   3rd Qu.: 0.670995  
##  Max.   : 17.5787   Max.   :1   Max.   : 1.1481061   Max.   : 4.024930  
##        x3                 eps1               eps2          
##  Min.   :-3.75367   Min.   :-5.86698   Min.   :-53.6324  
##  1st Qu.:-0.69583   1st Qu.:-0.95432   1st Qu.: -9.4102  
##  Median :-0.03440   Median : 0.03125   Median :  0.1831  
##  Mean   : 0.02018   Mean   : 0.01829   Mean   :  0.1411  
##  3rd Qu.: 0.66101   3rd Qu.: 0.97468   3rd Qu.:  9.8168  
##  Max.   : 3.77398   Max.   : 5.35372   Max.   : 59.9260
```

```
summary(DT)
```

```
##        y1                 const          x1                 x2          
##  Min.   :-11.7353   Min.   :1   Min.   :-1.0813748   Min.   :-3.626672  
##  1st Qu.: -0.2118   1st Qu.:1   1st Qu.:-0.2214267   1st Qu.:-0.686488  
##  Median :  2.2171   Median :1   Median : 0.0001322   Median : 0.000400  
##  Mean   :  2.1698   Mean   :1   Mean   :-0.0038033   Mean   :-0.006917  
##  3rd Qu.:  4.5833   3rd Qu.:1   3rd Qu.: 0.2136286   3rd Qu.: 0.670995  
##  Max.   : 17.5787   Max.   :1   Max.   : 1.1481061   Max.   : 4.024930  
##        x3                 eps1               eps2          
##  Min.   :-3.75367   Min.   :-5.86698   Min.   :-53.6324  
##  1st Qu.:-0.69583   1st Qu.:-0.95432   1st Qu.: -9.4102  
##  Median :-0.03440   Median : 0.03125   Median :  0.1831  
##  Mean   : 0.02018   Mean   : 0.01829   Mean   :  0.1411  
##  3rd Qu.: 0.66101   3rd Qu.: 0.97468   3rd Qu.:  9.8168  
##  Max.   : 3.77398   Max.   : 5.35372   Max.   : 59.9260
```

```
colnames(DT)
```

```
## [1] "y1"    "const" "x1"    "x2"    "x3"    "eps1"  "eps2"
```

```
#library(matlib)
X <- cbind(c(DT$const), c(DT$x1), c(DT$x2), c(DT$x3))

Y <-cbind(c(DT$y1))

XT <- t(X)

XTX <- XT%*%X
XTX
```

```
##             [,1]       [,2]       [,3]       [,4]
## [1,] 7584.00000  -28.84399  -52.45766 -153.04722
## [2,]  -28.84399  786.14496   10.09589  1794.58853
## [3,]  -52.45766   10.09589 7824.05912   79.67348
## [4,] -153.04722 1794.58853   79.67348 7616.46991
```

```
XTX_1 <- solve(XTX)
XTX_1
```

```
##               [,1]          [,2]          [,3]          [,4]
## [1,]  1.319182e-04 -2.660092e-06  8.546811e-07  3.254489e-06
## [2,] -2.660092e-06  2.752633e-03  3.036040e-06 -6.486578e-04
## [3,]  8.546811e-07  3.036040e-06  1.278334e-04 -2.035401e-06
## [4,]  3.254489e-06 -6.486578e-04 -2.035401e-06  2.842175e-04
```

```
XTY <- XT%*%Y
XTY
```

```
##          [,1]
## [1,] 16455.648
## [2,]  1174.928
## [3,] 10158.341
## [4,]  4738.788
```

```
beta_hat <- XTX_1%*%XTY
beta_hat
```

```
##           [,1]
## [1,] 2.1919799
## [2,] 0.1224008
## [3,] 1.3064804
## [4,] 0.6289678
```

This value is our OLS estimator

```
library(stargazer)
```

```
##
## Please cite as:
```

```
##  Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.
```

```
##  R package version 5.2.2. https://CRAN.R-project.org/package=stargazer
```

```
outl <- lm(y1 ~ x1 + x2 +x3, data=DT)
stargazer(outl, type = 'text')
```

```
##
## ===============================================
##                         Dependent variable:
##                     ---------------------------
##                                 y1
## -----------------------------------------------
## x1                             0.122
##                               (0.167)
##
## x2                            1.306***
##                               (0.036)
##
## x3                            0.629***
##                               (0.054)
##
## Constant                      2.192***
##                               (0.037)
##
## -----------------------------------------------
## Observations                   7,584
## R2                             0.179
## Adjusted R2                    0.179
## Residual Std. Error      3.183 (df = 7580)
## F Statistic          552.210*** (df = 3; 7580)
## ===============================================
## Note:               *p<0.1; **p<0.05; ***p<0.01
```

The parameter values from the matrix estimate are the same as from the regression estimate.

```
XTX/nrow(DT)
```

```
##              [,1]          [,2]         [,3]         [,4]
## [1,]  1.000000000 -0.003803269 -0.006916885 -0.02018028
## [2,] -0.003803269  0.103658354  0.001331209  0.23662823
## [3,] -0.006916885  0.001331209  1.031653365  0.01050547
## [4,] -0.020180277  0.236628234  0.010505469  1.00428137
```

```
cov(DT)
```

```
##               y1 const           x1           x2           x3         eps1
## y1    12.34255205     0  0.1631956969  1.354630521  0.67399025 -0.0345977698
## const  0.00000000     0  0.0000000000  0.000000000  0.00000000  0.0000000000
## x1     0.16319570     0  0.1036575573  0.001305074  0.23658268 -0.0009250523
## x2     1.35463052     0  0.0013050743  1.031741563  0.01036725  0.0044524846
## x3     0.67399025     0  0.2365826780  0.010367251  1.04840651 -0.0149200056
## eps1  -0.03459777     0 -0.0009250523  0.004452485 -0.01492001  2.0532128770
## eps2   0.35828556     0  0.0845584145 -0.105670109  0.22875506 -0.1667865058
##              eps2
## y1      0.35828556
## const   0.00000000
## x1      0.08455841
## x2     -0.10567011
## x3      0.22875506
## eps1   -0.16678651
## eps2  199.70843926
```

IV ESTIMATION

```
df.IV <- read.table("IV_Data.RData")
dt.IV <- data.table(df.IV)
colnames(dt.IV)
```

```
## [1] "y4"    "y5"    "const" "x1"    "x2"    "z1"    "z2"    "z3"    "z4"
```

```
pop.endog2 <- lm(y5 ~ x1 + x2, data=dt.IV)
stargazer(pop.endog2, type = "text")
```

```
##
## ===================================================
##                         Dependent variable:
##                     ---------------------------
##                                    y5
## ---------------------------------------------------
## x1                                2.930***
##                                   (0.007)
##
## x2                                1.680***
##                                   (0.007)
##
## Constant                          0.490***
##                                   (0.007)
##
## ---------------------------------------------------
## Observations                       17,584
## R2                                 0.932
## Adjusted R2                        0.932
## Residual Std. Error         0.965 (df = 17581)
## F Statistic           120,701.300*** (df = 2; 17581)
## ===================================================
## Note:                   *p<0.1; **p<0.05; ***p<0.01
```

```
cor.test(dt.IV$z4, dt.IV$x1)
```

```
##
##  Pearson's product-moment correlation
##
## data:  dt.IV$z4 and dt.IV$x1
## t = 2.2537, df = 17582, p-value = 0.02423
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.002214089 0.031766931
## sample estimates:
##       cor
## 0.01699422
```

2SLS

```
outlst <- lm(x1 ~ x2 + z3, data=dt.IV)
summary(outlst)
```

```
##
## Call:
## lm(formula = x1 ~ x2 + z3, data = dt.IV)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.8181 -0.6625  0.0070  0.6684  3.6077
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.002372   0.007369  -0.322    0.747
## x2           0.128994   0.007294  17.685   <2e-16 ***
## z3           0.186035   0.007397  25.150   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9771 on 17581 degrees of freedom
## Multiple R-squared:  0.05069,    Adjusted R-squared:  0.05059
## F-statistic: 469.4 on 2 and 17581 DF,  p-value: < 2.2e-16
```

```
dt.IV <- dt.IV[, x1hat:=predict(outlst, newdata=dt.IV)]
dt.IV

outJrd <- lm(y5 ~ x1hat + x2, data=dt.IV)
stargazer(outJrd, pop.endog2, type='text')
```

```
##
## ===================================================================
##                                   Dependent variable:
##                     ----------------------------------------------
##                                           y5
##                                (1)                    (2)
## -------------------------------------------------------------------
## x1hat                         2.115***
##                               (0.124)
##
## x1                                                  2.930***
##                                                     (0.007)
##
## x2                            1.784***              1.680***
##                               (0.028)               (0.007)
##
## Constant                      0.488***              0.490***
##                               (0.023)               (0.007)
##
## -------------------------------------------------------------------
## Observations                   17,584                17,584
## R2                             0.325                 0.932
## Adjusted R2                    0.325                 0.932
## Residual Std. Error (df = 17581)     3.044               0.965
## F Statistic (df = 2; 17581)     4,231.712***  120,701.300***
## ===================================================================
## Note:                                   *p<0.1; **p<0.05; ***p<0.01
```