

Econometrics_Problem_Set_6.R

Bernardo

Thu Oct 24 22:31:14 2019

```
library(readstata13)

## Warning: package 'readstata13' was built under R version 3.4.4
library(MASS)

kt <- read.dta13("kt_data.dta")

year10 <- subset(kt, year == 10)

firstreg <- lm(data = year10, wage ~ educ)
summary(firstreg)

##
## Call:
## lm(formula = wage ~ educ, data = year10)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.12602 -0.28743  0.03681  0.32398  1.42398
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.429075   0.081309   17.58  <2e-16 ***
## educ         0.073912   0.006206   11.91  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4939 on 1518 degrees of freedom
## Multiple R-squared:  0.08546,    Adjusted R-squared:  0.08486
## F-statistic: 141.9 on 1 and 1518 DF,  p-value: < 2.2e-16

secondreg <- lm(data = year10, wage ~ I(exper^2))
summary(secondreg)

##
## Call:
## lm(formula = wage ~ I(exper^2), data = year10)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.20571 -0.29571  0.02429  0.33855  1.40534
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.4387204  0.0275368  88.562  <2e-16 ***
## I(exper^2)  -0.0005309  0.0002416  -2.197   0.0282 *
## ---
```

```

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5156 on 1518 degrees of freedom
## Multiple R-squared:  0.00317,    Adjusted R-squared:  0.002513
## F-statistic: 4.828 on 1 and 1518 DF,  p-value: 0.02816

thirdreg <- lm(data = year10, wage ~ educ + I(exper^2) + abil + Moth_ed + Fath_ed)
summary(thirdreg)

##
## Call:
## lm(formula = wage ~ educ + I(exper^2) + abil + Moth_ed + Fath_ed,
##     data = year10)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.35714 -0.27677  0.04865  0.30607  1.35491
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.1498512  0.1298818   8.853  < 2e-16 ***
## educ         0.0746676  0.0084098   8.879  < 2e-16 ***
## I(exper^2)   0.0018881  0.0002811   6.716 2.63e-11 ***
## abil         0.0932260  0.0173086   5.386 8.34e-08 ***
## Moth_ed      -0.0043740  0.0059071  -0.740  0.4591
## Fath_ed       0.0104275  0.0046651   2.235  0.0255 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4816 on 1514 degrees of freedom
## Multiple R-squared:  0.1326, Adjusted R-squared:  0.1297
## F-statistic: 46.28 on 5 and 1514 DF,  p-value: < 2.2e-16

resid <- residuals(thirdreg)

year10 <- as.matrix(year10)

## calculate inverse matrix (X'x)^-1
X <- t(year10)
XX <- X %*% year10
inverse <- solve(XX)

minusierror <- as.numeric(vector())

## Calculate hii for each observation and calculate leave-one-out error
## through the formula e-i = ei(1-hii)^-1

for (i in 1:nrow(year10)){
  x_it <- t(as.numeric(year10[i, ]))
  x_i <- as.numeric(year10[i, ])
  hii <- x_it %*% inverse %*% x_i
  minusierror <- append(minusierror, resid[i]*(1-hii)^(-1))
}

plot(resid - minusierror)

```

