Advanced Quantitative Methods
Homework 3: Ordinary Least Squares

Johan A. Elkink
jos.elkink@ucd.ie

Due February 9, 2011, 9am

Please submit by email in PDF format. Add R code in a separate R file.

Overview of variables used below:

FEMALE	1 for females, 0 otherwise
REGUL_BUSINESS	Scale from 0 "Business and industry should be strictly regulated by the State" to 10 "Business and industry should be entirely free from regulation by the State"
IRELAND_IN_EU	Scale from 0 "Ireland's membership in the EU is a bad thing" to 10 "Ireland's membership of the EU is a good thing"
PUBLIC_VS_PRIVATE	Scale from 0 "Public or semi-state companies are the best way to provide the services people need" to 10 "Private enterprises are the best way of providing the services people need"
STATE_OWN	Scale from 0 "Most of business and industry should be owned by the state" to 10 "Most of business and industry should be privately owned"
HOMOSEXUALITY	Scale from 0 "Homosexuality is never justified" to 10 "Homosexuality is always justified"
UNIFICATION	Scale from 0 "European unification has already gone too far" to 10 "European unification should be pushed further"
GOD_EXISTS	Scale from 0 "God definitely does not exist" to 10 "God definitely does exist"
ENVIRONMENT	Scale from 0 "We should protect the environment even if this damages economic growth" to 10 "We should encourage economic growth even if this damages the environmnet"

1. Open the *ines_homework03.csv* data set and perform the following regression using matrix algebra instead of the built-in regression functions of R:

   $$UNIFICATION_i = \beta_0 + \beta_1 HOMOSEXUALITY_i + \beta_2 STATE\_OWN_i + \beta_3 GOD\_EXISTS_i + \varepsilon_i$$

   Note that for defining $\mathbf{y}$ and $\mathbf{X}$ you can use code along the lines of:

   ```
   ines <- read.csv(file.choose())
   y <- ines[,"UNIFICATION"]
   X <- ines[,c(1,"HOMOSEXUALITY","STATE_OWN","GOD_EXISTS")]
   ```

   (a) (10%) Report regression coefficients and standard errors

   (b) (10%) Plot the dependent variable against each independent variable, including partial regression lines

   (c) (10%) Plot the residuals against the dependent variable and all independent variables.

   (d) Based on just visual inspection of these plots, evaluate the extent to which the following assumptions hold for this model:

   - (5%) $E(\varepsilon|\mathbf{X}) = 0$
   - (5%) $\varepsilon \sim N(0, \sigma^2)$ (normally distributed errors)

- (5%) $var(\varepsilon|\mathbf{X}) = \sigma^2$ (homoscedastic)
- (5%) $cov(\varepsilon, \mathbf{X}) = 0$

(e) (10%) Calculate R^2 and adjusted R^2. How do you evaluate the overall fit of the model?

(f) (20%) Write a short (approximately 300 to 400 word) substantive interpretation of the results, which has to be consistent with the above results, but which should be phrased in layman terms, i.e. interpretable for anyone who does not understand regression analysis.

2. Using matrix algebra instead of built-in functions in R,

(a) (5%) Report regression coefficients and standard errors for: $STATE\_OWN_i = \beta_0 + \beta_1 FEMALE_i + \beta_2 ENVIRONMENT_i + \beta_3 GOD\_EXISTS_i + \varepsilon_i$

(b) (5%) Report regression coefficients and standard errors for: $STATE\_OWN_i = \beta_0 + \beta_1 FEMALE_i + u_i$

(c) (5%) Report regression coefficients and standard errors for regression the residuals of the previous regression on ENVIRONMENT and GOD_EXISTS, i.e. $u_i = \beta_0 + \beta_1 ENVIRONMENT_i + \beta_2 GOD\_EXISTS_i + v_i$

(d) (6%) What do you notice?